

# STATIONARY MEASURES AND EQUIDISTRIBUTION FOR ORBITS OF NON-ABELIAN SEMIGROUPS ON THE TORUS

JEAN BOURGAIN, ALEX FURMAN, ELON LINDENSTRAUSS, AND SHAHAR MOZES

ABSTRACT. Let  $\nu$  be a probability measure on  $\mathrm{SL}_d(\mathbb{Z})$  satisfying the moment condition  $\mathbb{E}_\nu(\|g\|^\epsilon) < \infty$  for some  $\epsilon$ . We show that if the group generated by the support of  $\nu$  is large enough, in particular if this group is Zariski dense in  $\mathrm{SL}_d$ , for any irrational  $x \in \mathbb{T}^d$  the probability measures  $\nu^{*n} * \delta_x$  tend to the uniform measure on  $\mathbb{T}^d$ . If in addition  $x$  is Diophantine generic, we show this convergence is exponentially fast.

## 1. INTRODUCTION AND STATEMENT OF THE MAIN RESULTS

Let  $\Gamma$  be a semigroup of  $d \times d$  nonsingular integer matrices, and consider the action of  $\Gamma$  on the torus  $\mathbb{T}^d$ . We assume throughout that the action is strongly irreducible: there is no subtorus invariant under a finite index subsemigroup of  $\Gamma$ .

The strong irreducibility assumption in particular implies that  $\Gamma$  acts ergodically on  $\mathbb{T}^d$  (equipped with the Lebesgue measure  $m$ ). Therefore the  $\Gamma$ -orbit of Lebesgue almost every  $x \in \mathbb{T}^d$  is dense, and in an appropriate sense even becomes equidistributed. However when  $\Gamma$  is cyclic there is a set of full Hausdorff dimension of exceptional points  $x$  for which  $\Gamma.x$  fails to be dense.

When  $\Gamma$  is bigger the distribution of individual  $\Gamma$ -orbits can be expected to be much more restrictive. An important result in this direction is due to Furstenberg who showed for  $d = 1$  (in which case  $\Gamma < \mathbb{Z}^\times$ , and in particular abelian) that if  $\Gamma$  is not virtually cyclic,  $\Gamma.x$  is dense for all irrational  $x \in \mathbb{T}$ , and moreover for any open  $U \subset \mathbb{T}$  there are only finitely many rational points whose  $\Gamma$ -orbits avoids  $U$ . This has been extended by Berend [1] to actions of abelian semigroups of toral endomorphisms on  $\mathbb{T}^d$ . However, in both cases, while the orbit closure of individual orbits are very restricted, there is some flexibility on how such an orbit distributes; for example consider the orbit of  $x = \sum_{k=1}^{\infty} 2^{-k!} \in \mathbb{T}$  under the semigroup  $\Gamma = \langle 2, 3 \rangle$ .

In this paper we consider the action of semigroups  $\Gamma$  which satisfy the following three conditions:

- ( $\Gamma$ -0)  $\Gamma < \mathrm{SL}_d(\mathbb{R})$ ,
- ( $\Gamma$ -1)  $\Gamma$  acts strongly irreducibly on  $\mathbb{R}^d$ ,

---

J.B. was supported in part by NSF grant DMS-0808042.

A.F. was supported in part by NSF grants DMS-0604611 and DMS-0905977.

E.L. was supported in part by NSF grants DMS-0554345 and DMS-0800345.

S.M. was supported in part by BSF and ISF.

( $\Gamma$ -2)  $\Gamma$  contain a proximal element: there is some  $g \in \Gamma$  with a dominant eigenvalue which is a simple root of its characteristic polynomial.

Note that ( $\Gamma$ -1) is substantially stronger than the requirement we have already imposed that  $\Gamma$  acts strongly irreducibly on  $\mathbb{T}^d$ . In particular, for  $d > 1$  an abelian semigroup never satisfies condition ( $\Gamma$ -1); indeed, the group generated by a semigroup satisfying ( $\Gamma$ -1) is nonamenable. Assumption ( $\Gamma$ -2) is a technical condition which is in particular satisfied when  $\Gamma$  is a Zariski dense semigroup of  $\mathrm{SL}_d(\mathbb{Z})$  [16]. While a substantial part of the argument works without assumption ( $\Gamma$ -0), without it simple counterexamples can be given to Theorem A below, similar to the example above of a non-equidistributed orbit for the semigroup  $\langle 2, 3 \rangle$ .

Under these (and more general) conditions, R. Muchnik [26] and Guivarc'h-Starkov [19] proved the analog of the theorems of Furstenberg and Berend, namely that for any  $x \in \mathbb{T}^d$  with at least one irrational coordinate  $\Gamma.x$  is dense, and moreover that there are only finitely many rational  $x$  whose orbit avoids a given open neighborhood in  $\mathbb{T}^d$ .

We study the quantitative distribution properties of  $\Gamma$ -orbits. Since  $\Gamma$  is not amenable we do this by considering a random walk on  $\Gamma.x$  corresponding to a probability measure  $\nu$  on  $\Gamma$ . We will assume that  $\nu$  satisfies the moment condition

$$(1.1) \quad \sum_{g \in \Gamma} \nu(g) \|g\|^\epsilon < \infty \quad \text{for some } \epsilon > 0.$$

Given a probability measure  $\nu$  on  $\Gamma$  and a probability measure  $\mu$  on  $\mathbb{T}^d$  the convolution  $\nu * \mu \in \mathbb{T}^d$  is

$$\nu * \mu = \sum_{g \in \Gamma} \nu(g) g_* \mu.$$

Furstenberg [14] has shown that under assumption ( $\Gamma$ -1) the top Liapunov exponent defined by

$$\lambda_1(\nu) = \lim_{n \rightarrow \infty} \frac{1}{n} \log \|g_1 g_2 \cdots g_n\| \quad \nu^{\mathbb{Z}_+}\text{-a.s.}$$

is positive. Assumption ( $\Gamma$ -2) guarantees that this Liapunov exponent is simple [16, 17]. Our main theorem is the following:

**Theorem A.** *Let  $\Gamma < \mathrm{SL}_d(\mathbb{R})$  satisfy ( $\Gamma$ -1) and ( $\Gamma$ -2) above, and let  $\nu$  be a probability measure supported on a set of generators of  $\Gamma$  satisfying (1.1). Then for any  $0 < \lambda < \lambda_1(\nu)$  there is a constant  $C = C(\nu, \lambda)$  so that if for a point  $x \in \mathbb{T}^d$  the measure  $\mu_n = \nu^{(n)} * \delta_x$  satisfies that for some  $a \in \mathbb{Z}^d \setminus \{0\}$*

$$|\widehat{\mu}_n(a)| > t > 0, \quad \text{with} \quad n > C \cdot \log\left(\frac{2\|a\|}{t}\right),$$

*then  $x$  admits a rational approximation  $p/q$  for  $p \in \mathbb{Z}^d$  and  $q \in \mathbb{Z}_+$  satisfying*

$$(1.2) \quad \left\| x - \frac{p}{q} \right\| < e^{-\lambda n} \quad \text{and} \quad |q| < \left( \frac{2\|a\|}{t} \right)^C.$$

This theorem has several corollaries:

**Corollary B.** *Let  $\Gamma$  and  $\nu$  be as in Theorem A, and  $x \in \mathbb{T}^d \setminus (\mathbb{Q}/\mathbb{Z})^d$ . Then the measures  $\mu_n = \nu^{*n} * \delta_x$  converge to the Haar measure  $m$  on  $\mathbb{T}^d$  in weak-\* topology.*

This answers affirmatively a question of Guivarc'h in a private communication, and should be contrasted with the example given above for the case of  $d = 1$ . We also have the following more quantitative equidistribution results:

**Corollary C.** *Let  $\Gamma$  and  $\nu$  be as in Theorem A,  $x \in \mathbb{T}^d$  and  $\mu_n = \nu^{*n} * \delta_x$ . Then there are  $c_1, c_2$  depending only on  $\nu$  so that the following holds:*

(1) *Assume  $x$  is Diophantine generic in the sense that for some  $M$  and  $Q$*

$$(1.3) \quad \left\| x - \frac{p}{q} \right\| > q^{-M} \quad \text{for all integers } q \geq Q \text{ and } p \in \mathbb{Z}^d.$$

*Then for  $n > c_1 \log Q$*

$$\max_{b \in \mathbb{Z}^d, 0 < \|b\| < B} |\widehat{\mu}_n(b)| < B e^{-c_2 n/M}$$

(2) *Assume  $x \notin (\mathbb{Q}/\mathbb{Z})^d$ . Then there is a sequence  $n_i \rightarrow \infty$  along which*

$$\max_{b \in \mathbb{Z}^d, 0 < \|b\| < e^{c_2 n_i}} |\widehat{\mu}_{n_i}(b)| < e^{-c_2 n_i}$$

Our next corollary answers a question raised by Furstenberg in [12]. Recall that a measure  $\mu$  is said to be  $\nu$ -stationary if  $\nu * \mu = \mu$ . If the support of  $\nu$  generates a semigroup  $\Gamma$ , any  $\Gamma$ -invariant probability measure is  $\nu$ -stationary for any probability measure  $\nu$  on  $\Gamma$ , but the converse (even for a fixed  $\nu$ ) is not true in general. Following Furstenberg ([12]), we say that an action  $\Gamma \curvearrowright X$  is  $\nu$ -stiff if any  $\nu$ -stationary measures is  $\Gamma$ -invariant.

In his paper [12] Furstenberg shows that for carefully chosen  $\nu$  on  $\mathrm{SL}_d(\mathbb{Z})$ , namely probability measures  $\nu$  so that the corresponding stationary measure on the boundary of  $\mathrm{SL}(d, \mathbb{R})$  is absolutely continuous with respect to Lebesgue, the action of  $\mathrm{SL}_d(\mathbb{Z})$  on  $\mathbb{T}^d$  is  $\nu$ -stiff. He then suggests that this should be true for any  $\nu$  whose support generates  $\mathrm{SL}_d(\mathbb{Z})$ . The following corollary of our main theorem confirms Furstenberg's insight:

**Theorem D.** *Let  $\Gamma < \mathrm{SL}_d(\mathbb{R})$  be a semigroup satisfying (Γ-1) and (Γ-2) above, and let  $\nu$  be a probability measure supported on a set of generators of  $\Gamma$  satisfying (1.1). Then any  $\nu$ -stationary measure  $\mu$  on  $\mathbb{T}^d$  is a convex combination of the Haar measure on  $\mathbb{T}^d$  and atomic measures supported by rational points. In particular, for such  $\nu$  the action of  $\Gamma$  on  $\mathbb{T}^d$  is  $\nu$ -stiff.*

The results of this paper have been announced in [6]. Since then an alternative, ergodic theoretic, approach to Theorem D was discovered by Y. Benoist and F. Quint [2]. This approach has the advantage of being more general; in particular, Benoist and Quint have been able to prove Theorem D without making the assumption (Γ-2). However their ergodic theoretic argument is not quantitative, certainly not in the sense of Theorem A. It also does not give equidistribution of  $\nu^{*n} * \delta_x$  as in Corollary B.

## 2. DEDUCTION OF COROLLARIES FROM THEOREM A

In this short section, we deduce Corollaries B and C from Theorem A. The deduction of Theorem D from Theorem A, or more precisely from the closely related Proposition 3.1, is given at the beginning of the next section.

*Proof of Corollary B given Theorem A.* Let  $x \in \mathbb{T}^d \setminus (\mathbb{Q}/\mathbb{Z})^d$ . Suppose that the measures  $\mu_n = \nu^{*n} * \delta_x$  fail to converge to the Haar measure  $m$ . Then by Weyl's equidistribution criterion it follows that for some  $a \in \mathbb{T}^d \setminus \{0\}$  and some sequence  $n_i \rightarrow \infty$

$$|\widehat{\mu}_{n_i}(a)| > t > 0 \quad \text{for all } i.$$

It follows from Theorem A that there is a sequence of approximations  $\frac{p_i}{q_i}$  tending to  $x$  with  $q_i$  uniformly bounded — which of course is only possible if  $x$  is rational.  $\square$

*Proof of Corollary C given Theorem A.* We first prove assertion (1) of the corollary. Let  $x$  be Diophantine generic in the sense of (1.3). Suppose that  $|\widehat{\mu}_n(b)| > t/B$  for some  $b \in \mathbb{Z}^d$  with  $0 < \|b\| < B$ . Then as long as

$$(2.1) \quad t > \frac{1}{2}e^{-n/C}$$

for  $C = C(\nu, \lambda_1/2)$  as in Theorem A, by (1.2) there are integers  $p, q$  so that

$$\left\| x - \frac{p}{q} \right\| < e^{-\lambda_1 n/2} \quad \text{and} \quad |q| < (2t^{-1})^C.$$

By (1.3), it follows that if  $n \geq c_1 \log Q$

$$(2.2) \quad e^{-\lambda_1 n/2} > q^{-M} > C't^{-MC}.$$

It follows from (2.1) and (2.2) that

$$t \leq C'' \max(e^{-n/C}, e^{-\lambda_1 n/2MC}),$$

establishing Corollary B, part(1).

Suppose now that for some  $x \notin (\mathbb{Q}/\mathbb{Z})^d$  part (2) of the corollary does not hold, i.e. that for every  $n$  there is a  $b_n \in \mathbb{Z}^d$  so that

$$|\widehat{\mu}_n(b_n)| \geq e^{-c_2 n}, \quad \text{and} \quad \|b_n\| < e^{c_2 n}.$$

Then by Theorem A, as long as  $2C_2 < 1$  and  $n$  is large enough there is a sequence of rational numbers  $\frac{p_n}{q_n}$  so that

$$(2.3) \quad \left\| x - \frac{p_n}{q_n} \right\| < e^{-\lambda_1 n/2} \quad \text{and} \quad |q_n| < 2^C e^{2c_2 C n}.$$

Since  $x$  is irrational, the sequence  $q_n$  is not eventually constant, so there are arbitrarily large  $n$  for which  $\frac{p_n}{q_n} \neq \frac{p_{n+1}}{q_{n+1}}$ . But then by (2.3) applied for both  $n, n+1$  gives

$$2^{-2C} e^{-4C_2(n+1)} \leq (q_n q_{n+1})^{-1} \leq \left\| \frac{p_n}{q_n} - \frac{p_{n+1}}{q_{n+1}} \right\| \leq 2e^{-\lambda_1 n/2},$$

which is a contradiction for large  $n$  if  $8C_2 < \lambda_1$ .  $\square$

## 3. OUTLINE OF THE PROOF

Given  $Q \in \mathbb{N}$  let

$$R_Q = \bigcup_{q \leq Q} \left\{ \left( \frac{p_1}{q}, \dots, \frac{p_d}{q} \right) \in \mathbb{T}^d : p_1, \dots, p_d \in \{0, \dots, q-1\} \right\}$$

denote the set of rational points on the torus with denominators  $q \leq Q$ . For  $r > 0$  let  $W_{Q,r} = \bigcup_{x \in R_Q} B_{x,r}$  denote the  $r$ -neighborhood of  $R_Q$ . We prove Theorem A by establishing the following:

**Proposition 3.1.** *Let  $\Gamma$  and  $\nu$  be as in Theorem A,  $0 < \lambda < \lambda_1(\nu)$ . Then for some constant  $C$  depending on  $\nu, \lambda$  the following holds: for any probability measure  $\mu_0$  on  $\mathbb{T}^d$ , if  $\mu_n = \nu^{*n} * \mu_0$  has a non-trivial Fourier coefficient  $a \in \mathbb{Z}^d \setminus \{0\}$*

$$(3.1) \quad |\widehat{\mu}_n(a)| > t, \quad \text{with } 0 < t < \frac{1}{2} \text{ and } n > C \cdot \log\left(\frac{2\|a\|}{t}\right),$$

then

$$(3.2) \quad \mu_0(W_{Q,e^{-\lambda n}}) > t^C \quad \text{where} \quad Q = \left(\frac{2\|a\|}{t}\right)^C.$$

By specializing to the case of  $\mu_0 = \delta_x$  we get Theorem A, since

$$\delta_x(W_{Q,e^{-\lambda n}}) > 0 \quad \iff \quad \left\| x - \frac{a}{q} \right\| < e^{-\lambda n} \quad \text{for some } q \leq Q.$$

Note that somewhat surprisingly Theorem A then implies a sharper form of Proposition 3.1 with the estimate (3.2) on the mass of almost rational points replaced by the sharper estimate  $\mu_0(W_{Q,e^{-\lambda n}}) > C't$ . In the special case of  $\mu = \mu_0 = \mu_1 = \dots$  a  $\nu$ -stationary probability measure, we can take  $n$  in Proposition 3.1 to be arbitrarily large, and deduce that for appropriate constant  $C$

$$\mu(R_Q) \geq t^C \quad \text{for } Q = \left(\frac{2\|a\|}{t}\right)^C,$$

with  $a$  and  $t$  as in (3.1), giving a somewhat more quantitative version of Theorem D.

We sketch the proof of Proposition 3.1. The proof consists of two phases:

- (Ph-1) first one starts with a lower bound on a single Fourier coefficient of the measure  $\mu_n = \nu^{*n} * \mu$ , namely  $|\widehat{\mu}_n(a)| > t$ , and deduce from this that for an appropriately chosen  $m_1 < n$  the measure  $\mu_{n-m_1}$  has a rich set of Fourier coefficients which are larger than a polynomial in  $t$ .
- (Ph-2) In the second phase, this information on the set of big Fourier coefficients of  $\mu_{n_1}$  for  $n_1 = n - m_1$  is used to show that for another appropriately chosen  $m_2 < n_1$  the measure  $\mu_{n_1-m_2}$  gives a significant (polynomial in  $t$ ) mass to small balls around rationals of low denominator.

It is perhaps instructive to present a proof of a much simpler result with a somewhat similar structure:

**Proposition 3.2** (“Baby Case”). *A probability measure  $\mu$  on  $\mathbb{T}^d$  which is  $\Gamma$ -invariant for  $\Gamma$  a finite index subgroup of  $\mathrm{SL}_d(\mathbb{Z})$  is a linear combination of Haar measure and a purely atomic  $\Gamma$ -invariant measure.*

In this setting one can use the following simple argument by Marc Burger [9].

- Assume that the  $\Gamma$ -invariant probability measure  $\mu$  is not Haar measure. Then  $\mu$  has a non trivial Fourier coefficient:

$$|\hat{\mu}(a)| = t > 0 \quad \text{at some } a \in \mathbb{Z}^d \setminus \{0\}.$$

Since  $\hat{\mu}(a) = \widehat{g_*\mu}(a) = \hat{\mu}(g^{\mathrm{tr}}a)$ , it follows that  $|\hat{\mu}(b)| = t_0 > 0$  for all  $b \in \Gamma^{\mathrm{tr}}a$ . For  $\mathrm{SL}_d(\mathbb{Z})$  and its finite index subgroups, any orbit  $\Gamma^{\mathrm{tr}}a \subset \mathbb{Z}^d \setminus \{0\}$  has *positive density* in  $\mathbb{Z}^d$ .

- By Wiener’s Lemma this implies that  $\mu$  has atoms. Indeed, evaluating  $\mu \times \mu(\Delta) = \mu * \check{\mu}(\{0\})$  (where  $\Delta$  is the diagonal in  $\mathbb{T}^d \times \mathbb{T}^d$  and the convolution  $\mu * \check{\mu}$  is the image of  $\mu \times \mu$  under the projection  $(x, y) \mapsto x - y$ ) in two ways one gets the identity

$$(3.3) \quad \sum_{x \text{ atom of } \mu} \mu(\{x\})^2 = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N \frac{1}{|B_n|} \sum_{b \in B_n} |\hat{\mu}(b)|^2$$

where  $B_N = \{a \in \mathbb{Z}^d : \max_{1 \leq i \leq d} |a_i| \leq N\}$ . It follows that any  $\gamma$ -invariant probability measure  $\mu$  on  $\mathbb{T}^d$  can be presented as a linear combination of Haar measure and a purely atomic  $\Gamma$ -invariant measure.

In the context of Proposition 3.1 establishing the existence of enough “big” Fourier coefficients for  $\mu_{n_1}$  given that  $\mu_n$  had at least one significant Fourier coefficient is substantially more involved, and we get much less than positive density. Consequently, in the second phase of the proof we will start with a weaker type of information on  $\mu$  than in the simple proof sketched above.

### 3.A. Phase I: Large scale structure of the set of large Fourier coefficients.

Starting from some  $a_0 \in \mathbb{Z}^d \setminus \{0\}$  with  $|\hat{\mu}_n(a_0)| = t_0 > 0$  for sufficiently large  $n$  depending on  $t_0, a_0$ , we shall prove that for  $t = t_0^p$  and any  $m_1$  in the range  $C(1 + \log t_0) < m_1 < n$  (with  $p, C$  some constants depending on  $\Gamma, \nu$ ) the set of  $t$ -“large” Fourier coefficients

$$(3.4) \quad A_{n-m_1, t} = \left\{ a \in \mathbb{Z}^d : |\hat{\mu}_{n_1}(a)| > t \right\}$$

is relatively “thick” in  $\mathbb{Z}^d$ , in the following sense.

Let  $\mathcal{N}(E; M)$  denotes the covering number of  $E \subset \mathbb{Z}^d$  by  $M$ -balls. In the simple proof of Proposition 3.2 the proportion of “large” Fourier coefficients in any sufficiently large box was shown to be positive. In the context of Proposition 3.1 the most difficult part of the proof, which in precise form is given by Theorem 6.1 below, gives that there is a large  $N$  (with  $\frac{N}{\|a_0\|}$  bounded above and below by an exponential in  $m_1$ ) and an exponentially smaller  $M$  (more precisely,  $\frac{M}{\|a_0\|}$  will be in

the range  $(\frac{N}{\|a_0\|})^{1-\kappa_1} < \frac{M}{\|a_0\|} < (\frac{N}{\|a_0\|})^{1-\kappa_2}$  so that the number of  $M$ -balls needed to cover the intersection  $A_{n_1,t} \cap [-N, N]^d$  is large — namely

$$(3.5) \quad \mathcal{N}\left(A_{n_1,t} \cap [-N, N]^d; M\right) > t_0^p (N/M)^d,$$

where  $p, \kappa_1, \kappa_2 > 0$  are constants depending only on  $\Gamma$  and  $\nu$ . Thinking of  $t_0$  as fixed (which is the case needed to establish Corollary B), this gives a lower bound on the covering number that is a positive proportion of the trivial upper bound.

To prove the key estimate (3.5), one starts with the identity

$$\widehat{\mu}_n(a_0) = \sum_g \nu^{*m}(g) \mu_{n-m}(g^{\text{tr}} a_0)$$

to conclude that if  $|\widehat{\mu}_n(a_0)| > t_0$ , then

$$(3.6) \quad \nu^{*m} \{g \in \Gamma : |\widehat{\mu}_{n-m}(g^{\text{tr}} a_0)| > t_0/2\} \geq t_0/2.$$

In Proposition 6.2 below we deduce from (3.6), using the quantitative theory of random matrix products, that once  $m_1$  is larger than some absolute constant,

$$(3.7) \quad \mathcal{N}\left(A_{n_1,t_1} \cap [-N_1, N_1]^d; M_1\right) > (N_1/M_1)^{\alpha_1};$$

where\*  $n_1 = n - m_1$ ,  $N_1 = \|a_0\| \exp(\frac{3}{2}\lambda m)$ ,  $M_1 = \|a_0\|$ ,  $t_1 = t_0/2$ , with  $\lambda$  the top Liapunov exponent corresponding to  $\nu$  (cf. Section 4).

For our proof it is crucial to improve the estimate (3.7) to the much sharper density type estimate (3.5). Equation (3.7) is equivalent to having an  $M_1$ -separated subset  $E \subset \mathbb{Z}^d \cap [-N_1, N_1]^d$  of cardinality  $|E| \geq (N_1/M_1)^{\alpha_1}$  so that for every  $a \in E$  we have  $|\widehat{\mu}_{n_1}(a)| > t_1$ ; and decreasing the cardinality of  $E$  by a constant factor we may assume

$$(3.8) \quad \left| \sum_{e \in E} \widehat{\mu}_{n_1}(e) \right| > t_1 |E|/2.$$

Similar to the way we used the identity  $\mu_n = \nu^{*m_1} * \mu_{n_1}$  in the proof of (3.7), equation (3.8) implies that (for any choice of  $m < n_1$ ), for  $\nu^{*m}$ -many  $g \in \Gamma$ , for many  $e \in g^{\text{tr}} E$  we have that  $|\widehat{\nu}_{n_1-m}(e)| > t_1/4$ ; indeed, if

$$\mathcal{G} = \{g \in \Gamma : |\{e \in g^{\text{tr}} E : |\widehat{\nu}_{n_1-m}(e)| > t_1/4\}| > t_1 |E|/4\}$$

then  $\nu^{*m}(\mathcal{G}) \geq t_1/4$ .

Our assumptions  $(\Gamma-0)$ – $(\Gamma-2)$  on  $\Gamma$  guarantee that the top Liapunov exponent for the random walk on  $\text{SL}_d(\mathbb{Z})$  corresponding to  $\nu$  is simple, which allows us to approximate  $\nu^{*m}$ -typical  $g$  by a composition of dilation (by a factor  $\sigma_1(g)$  in the range  $e^{(\lambda-\epsilon)m} \leq \sigma_1(g) \leq e^{(\lambda+\epsilon)m}$ ), a rotation, and a rank one projection, say  $\pi_g$ . The theory of random matrix products also gives us control over the distribution on the direction of the null space of this projection. Therefore choosing

---

\*There is nothing special about  $\frac{3}{2}$ ; any constant greater than one would do.

$M_2$  appropriately, we cannot distinguish with resolution  $M_2$  between the map  $a \mapsto g^{\text{tr}}a$  and this rank one transformation, e.g. in the sense that for any  $E' \subset E$

$$\mathcal{N}(g^{\text{tr}}(E'); M_2) \asymp \mathcal{N}(\sigma_1(g)\pi_g^{\text{tr}}(E'); M_2).$$

As long as  $m = m_2$  is sufficiently large (larger than some constant times  $|\log t|$ ), this applies to most  $g \in \mathcal{G}$  so that we can view  $g^{\text{tr}}(E)$  as a rotated and dilated rank one (random) projection of  $E$ .

If  $N_2, M_2, m_2$  are appropriately chosen, outside a set of  $g \in \Gamma$  of negligible  $\nu^{*m}$ -measure,  $g^{\text{tr}}([-N_1, N_1]^d)$  is contained in a rotated rectangular box of size  $[-N_2, N_2] \times [-M_2, M_2] \times \cdots \times [-M_2, M_2]$ . If  $\alpha_1$  were very close to  $d$  (say bigger than some  $\alpha_{\text{high}}$ ) we could use a variant on the Marstrand Projection Theorem, or more precisely on an extension due to Falconer [13], to show that for many  $g \in \mathcal{G}$

$$\mathcal{N}(g^{\text{tr}}E; M_2) \gg t_1^p(N_2/M_2),$$

and moreover that a similar inequality (with possibly a different implied constant, still polynomial in  $t_1$ ) holds for any subset  $E' \subset E$  with  $|E'| \geq t_1|E|/4$ . By definition of  $\mathcal{G}$ , one obtains that

$$\mathcal{N}\left(A_{n_2, t_1/4} \cap g^{\text{tr}}([-N_1, N_1]^d); M_2\right) \gg t_1^p(N_2/M_2),$$

and with some further arguments employing the inherent additive structure of Fourier coefficients of probability measures<sup>†</sup> get from this an estimate of the desired form

$$\mathcal{N}\left(A_{n_2, t_1/4} \cap [-N_2, N_2]^d; M_2\right) \gg t_1^{p'}(N_2/M_2)^d.$$

The argument sketched above is carried in Section 6.C below, and the resulting proposition is given by Proposition 6.5 below. Unfortunately, we have little control over  $\alpha_1$  which is determined by properties of the random walk corresponding to  $\nu$  on  $\text{SL}(d, \mathbb{R})$ . To handle the main case where  $\alpha_1 < \alpha_{\text{high}}$  we need to use arithmetic combinatorics: a projection result [5, Thm. 5] of the first author (based on techniques developed in the context of the Discretized Ring Conjecture [4]). Roughly stated, this theorem asserts that given a sufficiently rich set of lines  $D \subset \mathbb{P}^{d-1}$  and a (sufficiently non-degenerate) set  $E \subset [0, 1]^d$  of “dimension”  $\alpha$  there exist (many) lines  $\theta \in D$  so that the projection  $\pi_\theta(E)$  of  $E$  to  $\theta$  has “dimension”  $> (\alpha + \alpha_{\text{inc}})/d$ . This bootstrap step is the content of Proposition 6.3.

A complication in the proof of both Proposition 6.3 and Proposition 6.5 is that to employ the respective (discretized) projection theorem one needs finer control on the set to be projected than simply its covering number by  $M_i$ -balls. This is taken care of by zooming-in on a portion of the set  $A_{t_i, n_i} \cap [-N_i, N_i]$  in which there is greater regularity, and recentering this window using Cauchy-Schwartz, cf. Lemma 6.7.

<sup>†</sup>Essentially, the Cauchy-Schwartz inequality.



### 3.B. Phase II: Granulation structure of $\mu_0$ on the torus.

The information on the Fourier coefficients of a measure  $\mu_0$  for which a Fourier coefficient  $\widehat{\mu}_n(a)$  is significant that has been obtained in Phase I of the proof (with  $\mu_n = \nu^{*n} * \mu_0$  as before, and  $n$  sufficiently large depending on  $\|a\|$  and the size of  $|\widehat{\mu}_n(a)|$ ) can be translated to a statement about the measure  $\mu_0$  itself (and more generally about the measures  $\mu_{n-m}$  for  $m$  large enough) using the following elementary harmonic analysis proposition in the spirit of Wiener's Lemma:

**Proposition 3.3** (cf. Proposition 7.5). *If a probability measure  $\mu$  on  $\mathbb{T}^d$  satisfies (3.5) for some  $N > M$  then there exists a set  $X \subset \mathbb{T}^d$  of  $1/M$ -separated points in  $\mathbb{T}^d$  with*

$$\mu \left( \bigcup_{x \in X} \mathbf{B}_{x, \frac{1}{N}} \right) > t_0^{p'}.$$

Using this harmonic analytic fact, the outcome of the first stage of the proof is that for  $m_1 \gg \log(\|a\|/t_0)$ , the measure  $\mu_{n_1} = \mu_{n-m}$  is granulated in the following sense (cf. Proposition 7.1): for some constants  $1 < L_1 < L_2$  and  $\kappa > 0$  there is some  $\rho \in (L_2^{-m}, L_1^{-m})$  and the finite set  $X \subset \mathbb{T}^d$  so that

- (1)  $X$  is  $r = \rho^{1-\kappa}$ -separated;
- (2)  $\mu_{n_1} \left( \bigcup_{x \in X} \mathbf{B}_{x, \rho} \right) > t^C$ .

This is not yet what we want. So we continue with the strategy of successively sacrificing some convolution powers of  $\nu$  (i.e., increasing  $m$  to  $m' > m$ ) in exchange for more precise information on  $\mu_{n-m'}$ .

The two conditions (1)–(2) above on  $\mu_{n-m}$  and  $X$  guarantee in particular that  $t^{O(1)}$  of the mass of  $\mu_{n-m}$  is concentrated in balls of radius  $\rho$  whose measure is rather large, namely  $\geq t^{O(1)} \rho^{1-\kappa}$ .

Thanks to the separation condition, we can improve this estimate (cf. Proposition 7.2) and show that for appropriate  $m'$  (also  $\ll \log(\|a\|/t)$ ) there is a set  $X'$  of cardinality at most that of  $X$  so that  $\mu_{n-m'} \left( \bigcup_{x \in X'} \mathbf{B}_{x, \rho^N} \right) \geq t^{O_N(1)}$  for an arbitrary  $N$ .

At this stage we can rectify the unknown balls  $\{\mathbf{B}_{x, \rho^N} : x \in X'\}$  to be centered at rational points of controlled denominator. The reason for that is that as

$$\mu_{n-m'}(\mathbf{B}_{x, \rho^N}) = \sum_g \nu^{*\ell}(g) \mu_{n-m'-\ell}(g^{-1} \mathbf{B}_{x, \rho^N}),$$

if  $\mu_{n-m'}(\mathbf{B}_{x, \rho^N})$  is big, for many  $g$  with  $\|g^{-1}\|$  of controllable size (roughly  $e^{-\lambda_d \ell}$ , with  $\lambda_d$  the bottom Liapunov exponent of  $\nu$ ) the measure of the “shifted” balls  $\mu_{n-m'-\ell}(g^{-1} \mathbf{B}_{x, \rho^N})$  has to be big — so many  $g$  in fact that as  $\mu_{n-m'-\ell}$  is a probability measure there should be a lot of intersections between these shifted balls. These nontrivial intersections can be used to show that  $x$  is much closer to a rational of controlled denominator than what can be expected of a random point in  $\mathbb{T}^d$ . This rough scheme is carried out by Proposition 7.3.

Using the extra information obtained, one can proceed similarly to the first step mentioned above (i.e. Proposition 7.2) but with essentially no loss of mass (Proposition 7.4) and obtain the desired conclusion, i.e. Proposition 3.1.

## 4. RANDOM MATRIX PRODUCTS

**4.A. Notations.** Let  $G$  be a topological group; in this paper the discrete group  $\Gamma$  or the torus  $\mathbb{T}^d$ . On the set  $\text{Prob}(G)$  of all probability measures on  $G$  (for  $G = \mathbb{T}^d$  the measures are assumed to be Borel regular) one defines operations of *convolution*:  $\nu_1, \nu_2 \mapsto \nu_1 * \nu_2$ , and of a *reflection*  $\nu \mapsto \check{\nu}$ , by pushing forward  $\nu_1 \times \nu_2$  under the product map  $(g_1, g_2) \mapsto g_1 \cdot g_2$ , and pushing  $\nu$  by the inverse map  $g \mapsto g^{-1}$ , respectively. For  $n \in \mathbb{N}$  we write  $\nu^{*n}$  for the  $n$ th *convolution power* of  $\nu$  with itself. This should be distinguished from the product  $\nu^{\times n}$  defined on  $G^n$ .

Similarly, if  $G \curvearrowright X$  is a continuous action on a topological space, for  $\nu \in \text{Prob}(G)$  and  $\mu \in \text{Prob}(X)$  the *convolution*  $\nu * \mu \in \text{Prob}(X)$  is the pushforward of  $\nu \times \mu$  under the action map  $G \times X \rightarrow X$ . For  $\Gamma \curvearrowright \mathbb{T}^d$  and  $\nu \in \text{Prob}(\Gamma)$ ,  $\mu \in \text{Prob}(\mathbb{T}^d)$  we have

$$\nu * \mu = \sum_{g \in \Gamma} \nu(g) \cdot g_* \mu, \quad \text{where} \quad g_* \mu(E) = \mu(g^{-1}E).$$

For  $\mu \in \text{Prob}(\mathbb{T}^d)$  the Fourier coefficients are

$$\hat{\mu}(a) = \int_{\mathbb{T}^d} e_a(x) d\mu(x) \quad \text{where} \quad e_a(x) = e^{2\pi i \langle a, x \rangle} \quad (a \in \mathbb{Z}^d, x \in \mathbb{T}^d).$$

The Fourier transform intertwines  $\Gamma$ -actions on  $\mathbb{T}^d$  and on  $\mathbb{Z}^d = \hat{\mathbb{T}}^d$  according to

$$\widehat{g_* \mu}(a) = \hat{\mu}(g^{\text{tr}} a).$$

In a metric space (such as  $\mathbb{Z}^d, \mathbb{R}^d, \mathbb{P}^{d-1}, \mathbb{T}^d$ ) we denote by  $B_{x,r} = \{y : d(x,y) \leq r\}$  the closed  $r$ -ball around  $x$ , and by  $\text{Nbd}_r(E)$  the (closed)  $r$ -neighborhood of a set  $E$ .

For a set  $E$  denote by

$$\mathcal{N}(E; r) = \inf \left\{ n : \exists x_1, \dots, x_n \text{ s.t. } E \subset \bigcup_{i=1}^n B_{x_i, r} \right\}$$

the covering number of  $E$  by  $r$ -balls (these covering numbers will be used for finite subsets of  $\mathbb{Z}^d$  with a large  $r$ , and for subsets of  $\mathbb{P}^{d-1}$  and  $\mathbb{T}^d$  with small  $r > 0$ ).

**Linear algebra.** Throughout the paper we use the standard inner product  $\langle x, y \rangle = \sum_1^d x_i y_i$ , the Euclidean norm  $\|x\|^2 = \langle x, x \rangle$  on  $\mathbb{R}^d$ , and the operator norm  $\|g\| = \max \|gx\|/\|x\|$  on matrices  $g \in \text{GL}_d(\mathbb{R})$ . For  $x \in \mathbb{R}^d \setminus \{0\}$ ,  $\bar{x} \in \mathbb{R}x$  denotes the corresponding point in the projective space  $\mathbb{P}^{d-1}$ . We equip  $\mathbb{P}^{d-1}$  with the metric given by

$$d_{\text{ang}}(\bar{x}, \bar{y}) = \sin(\text{angle}(\bar{x}, \bar{y})) = \frac{\|x \wedge y\|}{\|x\| \cdot \|y\|}.$$

For  $g \in \mathrm{GL}_d(\mathbb{R})$  denote by  $\sigma_1(g) \geq \sigma_2(g) \geq \cdots \geq \sigma_d(g) > 0$  the *singular values* of  $g$ . In the polar decomposition we have

$$g = U \begin{pmatrix} \sigma_1(g) & & \\ & \ddots & \\ & & \sigma_d(g) \end{pmatrix} V \quad \text{with} \quad U, V \quad \text{orthogonal.}$$

For  $g \in \mathrm{GL}_d(\mathbb{R})$  let  $\rho(g) = \sigma_2(g)/\sigma_1(g)$ . If  $\rho(g) < 1$  let

$$\theta(g) = U\bar{e}_1 \in \mathbb{P}^{d-1}.$$

This is the direction of the long axis of the  $g$  image of the round ball  $\{x \in \mathbb{R}^d : \|x\| \leq 1\}$ . Denote by  $H(g)$  the hyperplane of vectors with “shorter stretch”

$$\begin{aligned} H(g) &= \left\{ \bar{z} \in \mathbb{P}^{d-1} : Vz \in \mathrm{Span}(e_2 + \cdots + e_d) \right\} \\ &\subset \left\{ \bar{z} \in \mathbb{P}^{d-1} : \|gz\| \leq \sigma_2(g)\|z\| \right\} \end{aligned}$$

Note that  $\theta(g)$  describes the direction of the *image* of the “long vector”, under  $g : \mathbb{R}^d \rightarrow \mathbb{R}^d$ , while  $H(g)$  refers to the *source* of the shorter ones.

**Lemma 4.1.** *For  $g \in \mathrm{GL}_d(\mathbb{R})$  with  $\rho(g) < 1$ .*

(1)  $H(g) = \theta(g^{\mathrm{tr}})^\perp$ .

(2) For any  $0 \neq z \in \mathbb{R}^d$ ,

$$\|g\| \cdot \|z\| \cdot d_{\mathrm{ang}}(\bar{z}, H(g)) \leq \|gz\| \leq \|g\| \cdot \|z\| \cdot (\rho(g) + d_{\mathrm{ang}}(\bar{z}, H(g))).$$

(3)  $d_{\mathrm{ang}}(g\bar{z}, \theta(g)) < \rho(g)/d_{\mathrm{ang}}(\bar{z}, H(g))$  for any  $0 \neq z \in \mathbb{R}^d$ .

(4) If  $g = hk$  with  $\rho(g) < 1$  and  $2\rho(h) < \|g\|/(\|h\| \cdot \|k\|)$  then

$$d_{\mathrm{ang}}(\theta(g), \theta(h)) < 2\rho(h) \cdot \frac{\|h\| \cdot \|k\|}{\|g\|}.$$

*Proof.* (1) is immediate from the definitions.

(2) Write  $z = \|z\| \cdot (tx + sy)$  with  $\bar{x} \in H(g)^\perp$ ,  $\bar{y} \in H(g)$ ,  $\|x\| = \|y\| = 1$ . Then  $|t| = d_{\mathrm{ang}}(\bar{z}, H(g))$ , while

$$\|z\| \cdot \|g\| \cdot |t| \leq \|gz\| = \|z\| \cdot \sqrt{t^2\|gx\|^2 + s^2\|gy\|^2} \leq \|z\| \cdot (|t|\|g\| + |s|\sigma_2(g)).$$

(3) Assume  $\|z\| = 1$  and write  $z = tx + sy$  as in (2). We have  $\theta(g) = g\bar{x}$  and  $\|g\| = \|gx\|$  and  $\|gz\| \geq \|g\| \cdot |t|$ . Also  $gx \wedge gz = gx \wedge (tgx + sgy) = s(gx \wedge gy)$ . Hence

$$d_{\mathrm{ang}}(g\bar{z}, \theta(g)) = \frac{\|gz \wedge gx\|}{\|gz\| \cdot \|gx\|} \leq \frac{|s| \cdot \|gy\| \cdot \|gx\|}{\|g\| \cdot |t| \cdot \|gx\|} \leq \frac{\|gy\|}{\|g\| \cdot |t|}.$$

Now (3) follows, because  $\|gy\| \leq \sigma_2(g)$  and  $|t| = d_{\mathrm{ang}}(\bar{z}, H(g))$ .

(4) Choose a unit vector  $x \perp H(g)$ , denote  $z = kx$  and write

$$z = \|z\| \cdot (ty + sw) \quad \text{with} \quad y \in H(h)^\perp, \quad w \in H(h), \quad \|y\| = \|w\| = 1.$$

Thus  $d_{\text{ang}}(\bar{z}, H(h)) = |t|$ . We have

$$\|g\| = \|gx\| = \|hz\|, \quad \|z\| = \|kx\| \leq \|k\| \implies \frac{\|hz\|}{\|z\|} \geq \frac{\|g\|}{\|k\|}.$$

But  $\|hz\|^2 \leq \|z\|^2(t^2\sigma_1(h)^2 + \sigma_2(h)^2s^2)$  because  $w \in H(h)$ . Hence

$$\frac{\|g\|}{\|h\| \cdot \|k\|} \leq \frac{\|hz\|}{\|h\| \cdot \|z\|} \leq \sqrt{t^2 + \rho(h)^2s^2} \leq \sqrt{t^2 + \rho(h)^2}.$$

Denoting by  $c$  the LHS, we get  $d_{\text{ang}}(\bar{z}, H(h)) = |t| \geq \sqrt{c^2 - \rho(h)^2}$ . Since  $\theta(g) = g\bar{x} = h\bar{z}$ , estimate (3) gives

$$d_{\text{ang}}(\theta(g), \theta(h)) = d_{\text{ang}}(h\bar{z}, \theta(h)) \leq \frac{\rho(h)}{\sqrt{c^2 - \rho(h)^2}} < \frac{2\rho(h)}{c}$$

under the assumption  $2\rho(h) < c$ .  $\square$

**4.B. Random walks.** Let  $\nu$  be a probability measure on  $\text{SL}_d(\mathbb{R})$  such that

$$(4.1) \quad \int \log \|g\| d\nu < \infty$$

The Lyapunov exponents  $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_d$  of  $\nu$  are defined through the limits of the following sub-additive sequences:

$$\lambda_1 = \lim_{n \rightarrow \infty} \int \frac{1}{n} \log \|g\| d\nu^{*n}(g), \quad \sum_{i=1}^k \lambda_i = \lim_{n \rightarrow \infty} \int \frac{1}{n} \log \|\wedge^k g\| d\nu^{*n}(g).$$

Equivalently,  $\lambda_i$  describes the asymptotic of  $\int n^{-1} \cdot \log \sigma_i(g) d\nu^{*n}(g)$ , where  $\sigma_i$  are the singular values; in particular,  $\sigma_1(g) = \|g\|$ . The convergence holds not only on average, but also a.e. and in  $L^1$ : if  $(g_1, g_2, \dots)$  are chosen independently according to  $\nu$  then, using Kingman's subadditive ergodic theorem, with probability one and in  $L^1(\nu^\infty)$  a long random product has polar decomposition

$$g_n \cdots g_2 g_1 = U \begin{pmatrix} e^{\lambda_1 n + o(n)} & & \\ & e^{\lambda_2 n + o(n)} & \\ & & \ddots \end{pmatrix} V$$

with  $U$  and  $V$  orthogonal.

**Theorem 4.2** ([17], [16]). *Let  $\nu$  be a probability measure on  $\text{SL}_d(\mathbb{R})$  with (4.1) and so that the group  $\langle \text{supp}(\nu) \rangle$  satisfies conditions  $(\Gamma-0)$ – $(\Gamma-2)$  of p. 1. Then the top Lyapunov exponent is simple:*

$$\lambda_1 > \lambda_2$$

*In particular,  $\lambda_1 > 0$ .*

If  $\langle \text{supp}(\nu) \rangle$  is irreducible on  $\mathbb{R}^d$  then ([15]) for any fixed  $x \in \mathbb{R}^d \setminus \{0\}$  for  $\nu^\infty$ -a.e. sequence  $(g_1, g_2, \dots)$

$$\frac{1}{n} \log \|g_n \cdots g_1 x\| = \lambda_1$$

and, denoting  $h_n = g_n \cdots g_1$ , the angular distance  $d_{\text{ang}}(h_n \bar{x}, \theta(h_n)) \rightarrow 0$ .

We shall need exponential estimates for various rates of convergence in the above stated limits. Such estimates are known under an assumption slightly stronger than (4.1), namely:

$$(4.2) \quad \int \|g\|^\epsilon d\nu(g) < \infty \quad \text{for some } \epsilon > 0.$$

**Theorem 4.3** (Large deviations). *Let  $\nu \in \text{Prob}(\text{SL}_d(\mathbb{R}))$  satisfy (4.2). Then for any  $\omega > 0$  there is  $\rho_\omega > 0$  and  $m_0(\omega)$  so that for  $m \geq m_0(\omega)$*

$$\begin{aligned} \nu^{*m} \left\{ g : \left| \lambda_1 - \frac{1}{m} \log(\|gx\| / \|x\|) \right| > \omega \right\} &< e^{-\rho_\omega \cdot m} \quad \forall x \in \mathbb{R}^{d-1} \setminus \{0\} \\ \nu^{*m} \left\{ g : \left| \lambda_i - \frac{1}{m} \log \sigma_i(g) \right| > \omega \right\} &< e^{-\rho_\omega \cdot m} \quad (i = 1, \dots, d) \end{aligned}$$

*Proof.* The first inequality follows from [3, Thm. V.6.1] and the remarks following the proof regarding uniformity in  $x$ ; the second inequality is [3, Thm. V.6.2].  $\square$

**Theorem 4.4** (Exponential Estimates). *Let  $\nu \in \text{Prob}(\text{SL}_d(\mathbb{R}))$  satisfy (4.2) and conditions  $(\Gamma-0)$ – $(\Gamma-2)$  of p. 1. Then for any  $c_1 > 0$  and some  $c_2 > 0$  there exist  $c_3 > 0$  and  $m_0 \in \mathbb{N}$  so that for all  $\bar{x}, \bar{y} \in \mathbb{P}^{d-1}$  each of the following subsets of  $\Gamma$ :*

- (1)  $\{g \in \Gamma : d_{\text{ang}}(g\bar{x}, \bar{y}) > e^{-c_1 \cdot m}\},$
- (2)  $\{g \in \Gamma : d_{\text{ang}}(g\bar{x}, \bar{y}^\perp) > e^{-c_1 \cdot m}\}$
- (3)  $\{g \in \Gamma : d_{\text{ang}}(g\bar{x}, \theta(g)) < e^{-c_2 \cdot m}\}$

has  $\nu^{*m}$ -probability  $> 1 - e^{-c_3 \cdot m}$  for  $m \geq m_0$ .

*Proof.* We first establish (3). By Theorem 4.3, there is some  $\rho_1 > 0$  so that with probability  $> 1 - e^{-\rho_1 m}$ ,

$$(4.3) \quad \begin{aligned} \left| \frac{1}{m} \log \sigma_1(g) - \lambda_1 \right| &< \frac{\lambda_1 - \lambda_2}{12} \\ \left| \frac{1}{m} \log \sigma_2(g) - \lambda_2 \right| &< \frac{\lambda_1 - \lambda_2}{12}. \end{aligned}$$

In the notations of Lemma 4.1, the above two inequality implies that  $\rho(g) < e^{-5(\lambda_1 - \lambda_2)m/6}$ , hence by (3) of that lemma, at least one of the following two equations holds:

$$(4.4) \quad d_{\text{ang}}(g\bar{x}, \theta(g)) < e^{-(\lambda_1 - \lambda_2)m/2}$$

$$(4.5) \quad d_{\text{ang}}(\bar{x}, H(g)) < e^{-(\lambda_1 - \lambda_2)m/3}.$$

But if (4.5) holds, by Lemma 4.1.(2)

$$(4.6) \quad \begin{aligned} \log(\|gx\| / \|x\|) &\leq 1 + \max(\log \sigma_1(g) - m(\lambda_1 - \lambda_2)/3, \log \sigma_2(g)) \\ &\leq 1 + (2\lambda_1 + \lambda_2)/3 \end{aligned}$$

By Theorem 4.3, (4.6) occurs with probability  $< e^{-\rho_2 m}$  for some positive  $\rho_2$  and  $m > m_0$ , so  $d_{\text{ang}}(g\bar{x}, \theta(g)) < e^{-(\lambda_1 - \lambda_2)m/2}$  outside a set of  $\nu^{*m}$ -measure  $\leq e^{-\rho_1 m} + e^{-\rho_2 m}$ , establishing (3) with  $c_2 = (\lambda_1 - \lambda_2)/2$ .

We now turn to the proof of assertion (2). This also relies on Theorem 4.3, but applied to the random walk corresponding to the measure  $\tilde{\nu}$  defined by  $\tilde{\nu}(g) = \nu(g^{\text{tr}})$ . The measure  $\tilde{\nu}$  also satisfies conditions (4.2) and conditions  $(\Gamma-0)$ – $(\Gamma-2)$  above and moreover has the same the Lyapunov exponents as  $\nu$ . Writing

$$g = U \begin{pmatrix} \sigma_1(g) & & \\ & \ddots & \\ & & \sigma_d(g) \end{pmatrix} V \quad \text{with} \quad U, V \quad \text{orthogonal.}$$

and recalling that by definition  $\theta(g) = Ue_1$ , we have that if (4.4) holds

$$\begin{aligned} d_{\text{ang}}(gx, y^\perp) &\geq d_{\text{ang}}(\theta(g), y^\perp) - d_{\text{ang}}(gx, \theta(g)) \\ &\geq d_{\text{ang}}(e_1, (U^{\text{tr}}y)^\perp) - e^{-(\lambda_1 - \lambda_2)m/2} \\ &\geq \frac{\|g^{\text{tr}}y\| / \|y\| - \sigma_2(g)}{\sigma_1(g)} - e^{-(\lambda_1 - \lambda_2)m/2}. \end{aligned}$$

By Theorem 4.3, with probability  $\geq 1 - e^{-\rho_3 m}$

$$\frac{\|g^{\text{tr}}y\| / \|y\| - \sigma_2(g)}{\sigma_1(g)} \geq 2e^{-\min(c_1, (\lambda_1 - \lambda_2)/2)m},$$

hence  $d_{\text{ang}}(gx, y^\perp) \geq 2e^{-c_1 m}$  with probability  $\geq 1 - e^{-\rho_1 m} - e^{-\rho_2 m} - e^{-\rho_3 m}$ , establishing (2).

Assertion (1) is a trivial consequence of (2).  $\square$

**4.C. Some further estimates.** In this subsection we shall establish some basic estimates that will be used in the following sections.

Given a point  $\bar{x}$  and a set  $D$  in  $\mathbb{P}^{d-1}$  let

$$\Gamma_{\bar{x} \rightarrow D} = \{g \in \Gamma : g\bar{x} \in D\}, \quad \Gamma_{\theta(g) \in D} = \{g \in \Gamma : \theta(g) \in D\}.$$

**Lemma 4.5** (Basic estimate of distribution of directions). *There exist  $\tau > 0$  and  $m_0$  so that for any  $r$  in the range*

$$e^{-m} < r < e^{-m_0},$$

one has

$$\nu^{*m}(\Gamma_{\bar{x} \rightarrow V}) < r^\tau, \quad \nu^{*m}(\Gamma_{\theta(g) \in V}) < r^\tau.$$

for any  $\bar{x} \in \mathbb{P}^{d-1}$  and an  $r$ -neighborhood of a hyperplane  $V = \text{Nbd}_r(\bar{y}^\perp)$ .

*Proof.* Given  $r > 0$  let  $V(r) = \text{Nbd}_r(\bar{y}^\perp)$ , and choose  $k \in \mathbb{N}$  with

$$e^{-k-1} < r \leq e^{-k}.$$

Assuming  $k > m_0$  for  $m_0$  as in Theorem 4.4 applied with  $c_1 = 1$ , by part (2) of that theorem

$$\begin{aligned} \nu^{*m}(\Gamma_{\bar{x} \rightarrow V(r)}) &\leq \nu^{*m}(\Gamma_{\bar{x} \rightarrow V(e^{-k})}) = \int \nu^{*k}(\Gamma_{h\bar{x} \rightarrow V(e^{-k})}) d\nu^{*(m-k)}(h) \\ &\leq \int e^{-c_2 k} d\nu^{*(m-k)}(h) = e^{-c_5 \cdot k} \\ &< (er)^{c_3}, \end{aligned}$$

with  $c_3$  as in Theorem 4.4.

To estimate  $\nu^{*m}(\Gamma_{\theta(g) \in V(r)})$  set  $r_1 = \max(r, e^{-c_2 m})$ . Then

$$\nu^{*m}(\Gamma_{\theta(g) \in V(r)}) \leq \nu^{*m}(\Gamma_{\theta(g) \in V(2r_1)}) + \nu^{*m}(d_{\text{ang}}(g\bar{x}, \theta(g))).$$

The first of these is  $\leq (2er_1)^{c_3}$ ; the second, by part (3) of Theorem 4.4, is  $\leq r_1^{c_3}$ . It follows that

$$\nu^{*m}(\Gamma_{\theta(g) \in V(r)}) \leq 2(2er_1)^{c_3} \leq Cr^{(\min(1, c_2)c_3)}.$$

Setting e.g.  $\tau = \min(c_3, c_2 c_3)/2$  the lemma follows once  $m_0$  is sufficiently large.  $\square$

Given a set  $F = \{\bar{x}_1, \dots, \bar{x}_d\} \subset \mathbb{P}^{d-1}$  a quantitative measure of the extent to which these lines are in general position is given by the volume spanned by unit vectors in these directions:

$$\text{vol}(\bar{x}_1, \dots, \bar{x}_d) = \frac{|x_1 \wedge \dots \wedge x_d|}{\|x_1\| \dots \|x_d\|}.$$

This quantity is symmetric in the arguments, but can be computed as

$$\text{vol}(\bar{x}_1, \dots, \bar{x}_d) = \prod_{i=2}^d d_{\text{ang}}(x_i, \text{Span}(x_1, \dots, x_{i-1})).$$

Hence, denoting

$$u(\bar{x}_1, \dots, \bar{x}_d) = \min_{1 \leq j \leq d} d_{\text{ang}}(x_j, \text{Span}(x_1, \dots, \hat{x}_j, \dots, x_d))$$

we have

$$u(\bar{x}_1, \dots, \bar{x}_d)^d \leq \text{vol}(\bar{x}_1, \dots, \bar{x}_d) \leq u(\bar{x}_1, \dots, \bar{x}_d).$$

**Lemma 4.6** (General position). *For some  $p < \infty, c_0$  and  $s_0 > 0$  depending on  $\nu$ , one has*

$$(\nu^{*m})^{\times d} \left\{ \vec{g} \in \Gamma^d : \text{vol}(\theta(g_1), \dots, \theta(g_d)) > s^p \right\} > 1 - s.$$

and

$$(\nu^{*m})^{\times d} \left\{ \vec{g} \in \Gamma^d : \text{vol}(\theta(g_1^{\text{tr}}), \dots, \theta(g_d^{\text{tr}})) > s^p \right\} > 1 - s.$$

for  $e^{-cm} < s < s_0$

*Proof.* Let  $r = d^{-1} \cdot s^{1/\tau}$ . Given any arbitrary  $g_1 \in \Gamma$  the  $\nu^{*m}$ -probability that

$$d_{\text{ang}}(\theta(h), \theta(g_1)) > r$$

is at least  $1 - r^\tau$  (Theorem 4.4). For the same reason given any  $g_1, g_2$ :

$$\nu^{*m} \{ h : d_{\text{ang}}(\theta(h), \theta(g_1) \oplus \theta(g_2)) > r \} > 1 - r^\tau.$$

Continuing this argument we deduce that the set

$$\left\{ \vec{g} \in \Gamma^d : d_{\text{ang}}(\theta(g_i), \theta(g_1) \oplus \cdots \oplus \theta(g_{i-1})) > r, i = 2, \dots, d \right\}$$

has  $(\nu^{*m})^{\times d}$ -measure at least

$$(1 - r^\tau)^{d-1} > 1 - (d-1)r^\tau > 1 - s.$$

On the other hand every  $d$ -tuple in the set above has

$$\text{vol}(\theta(g_1), \dots, \theta(g_d)) > r^d$$

If  $p$  is large enough,  $s^p < r^d = d^{-d}s^{d/\tau}$ .

To deduce the second estimate, one may apply the same arguments to the random walk generated by  $\tilde{\nu}$ , with  $\tilde{\nu}$  the transpose to  $\nu$  as in p. 14.  $\square$

## 5. TWO NOTIONS OF COARSE DIMENSION

Given a subset  $\tilde{A}$  of  $\mathbf{B}_{0,1} \subset \mathbb{R}^d$ , there are several ways one can try to estimate its dimension, or more precisely, in our case, its dimension at scale  $r$ . One simple way is via covering numbers: we can consider  $\tilde{A}$  to be of ‘‘coarse dimension’’  $\geq \alpha$  at scale  $r$  if  $\mathcal{N}(\tilde{A}; r) \geq r^{-\alpha}$ . Another, more restrictive definition of ‘‘coarse dimension  $\geq \alpha$ ’’ is via the following:

**Definition 5.1.** A measure  $\rho$  on a set  $B$  is said to be  $(C, \alpha)$ -regular at scale  $r$  on  $B$  if for any  $x \in A$ ,  $s \geq r$

$$\rho(\mathbf{B}_{x,s}) < C \left( \frac{s}{\text{diam } B} \right)^\alpha.$$

Thus another plausible definition of ‘‘coarse dimension’’ of a finite set  $A$  would be that  $A$  supports some probability measure  $\rho$  which is  $(C, \alpha)$ -regular at scale  $r$  on  $A$  for some absolute constant  $C$ .

The following lemma allows us to relate the two notions:

**Lemma 5.2.** *For any  $\epsilon > 0$  there are constants  $C_\epsilon, C'_\epsilon > 0$  such that for every  $s, \alpha$  with  $2\epsilon < s < \alpha$  and  $r < 1$ , if  $\tilde{A} \subset \mathbf{B}_{0,1} \subset \mathbb{R}^d$  satisfies*

$$\mathcal{N}(\tilde{A}; r) \geq r^\alpha$$

*then there is a point  $x \in \mathbf{B}_{0,1}$  and a probability measure  $\rho$  supported on  $\tilde{A} \cap \mathbf{B}_{x,r,\beta}$  which is  $(C_\epsilon, \alpha - s)$ -regular on  $\mathbf{B}_{x,C'_\epsilon r^\beta}$  at scale  $r$  for  $\beta = \frac{d-\alpha+\epsilon}{d-\alpha+s-\epsilon}$ .*

*Proof.* Let  $T$  be a large integer (which will eventually be determined by  $\epsilon$ ), and  $k_1 = \lceil -\log_2(r)/T \rceil$ . Without loss of generality we shall assume that every cube of size  $2^{-k_1 T}$  intersects  $\tilde{A}$  in at most one point.

Consider

$$\mathcal{Q} = \left\{ (x_1, \dots, x_d) : \exists 1 \leq i \leq d, \exists 0 \leq k \leq k_1 \text{ so that } 2^{kT} x_i \in \mathbb{Z} + [0, 2^{-T}] \right\}.$$

Then as  $\mathcal{Q}$  is a subset of  $\mathbb{R}^d$  of density at most

$$1 - (1 - d2^{-T})^{k_1}$$



there is a translate  $\tilde{A} + \xi$  of  $\tilde{A}$  so that

$$\left| (\tilde{A} + \xi) \setminus \mathcal{Q} \right| \geq (1 - d2^{-T})^{k_1} \left| \tilde{A} \right| \geq C^{(1)} r^{-\alpha + \epsilon/2}$$

as long as  $T$  is large enough (depending only on  $d, \epsilon$ ) for some constant  $C^{(1)}$  (depending on  $d, T$  and  $\epsilon$ ). Let  $\tilde{A}_0 = (\tilde{A} + \xi) \setminus \mathcal{Q}$ .

We shall call a cube of the form  $[n_1 2^{-kT}, (n_1 + 1) 2^{-kT}) \times \cdots \times [n_d 2^{-kT}, (n_d + 1) 2^{-kT})$  for  $(n_1, \dots, n_d) \in \mathbb{Z}^d$  a  $2^{-kT}$ -cube. By definition of  $\tilde{A}_0$ , for any  $0 \leq k \leq k_1$ , and any two distinct  $2^{-kT}$ -cubes  $I_1, I_2$  intersecting  $\tilde{A}_0$ , the distance between  $I_1 \cap \tilde{A}_0$  and  $I_2 \cap \tilde{A}_0$  is at least  $2^{-(k+1)T}$  (this is precisely the purpose of removing points of  $\mathcal{Q}$  from an appropriate shift of  $\tilde{A}$ ).

It will be convenient to start by extracting from  $\tilde{A}_0$  a large subset  $\tilde{A}_1$  with tree-structure (similar to but simpler than that used in [4, 5]). By this we mean that there are integers  $R_1, \dots, R_{k_1}$  with  $1 \leq R_k \leq 2^T$  so that if  $\mathcal{A}_k$  denotes the collection of  $2^{-kT}$ -cubes intersecting  $\tilde{A}_1$ , then each  $0 \leq k < k_1$ , each  $2^{-kT}$ -cube  $I \in \mathcal{A}_k$  contains precisely  $R_{k+1}$  cubes in  $\mathcal{A}_{k+1}$ . By successively trimming the set  $\tilde{A}_0$  one easily shows that if  $T$  is large enough (also depending only on  $\epsilon$ ), one can find such a subset  $\tilde{A}_1 \subset \tilde{A}_0$  with tree-structure so that  $\left| \tilde{A}_1 \right| \geq C^{(2)} r^{-\alpha + \epsilon}$ .

Indeed, to obtain this trimmed set  $\tilde{A}_1$ , consider all  $2^{-(k_1-1)T}$  cubes intersecting  $\tilde{A}_0$ , and find  $R_{k_1}$  so that the number of these cubes containing between  $R_{k_1}$  and  $2R_{k_1}$  of the  $2^{-k_1 T}$ -cubes is maximized. Throw away all points of  $\tilde{A}_0$  which are not contained in such a  $2^{-(k_1-1)T}$ -cube. Suppose  $Q$  is one of the remaining  $2^{-(k_1-1)T}$ -cubes, and that exactly  $n_Q$  of the  $2^{k_1 T}$ -subcubes in  $Q$  have nonempty intersection with  $\tilde{A}_0$ . We throw away all points of  $\tilde{A}_0$  in  $n_Q - R_{k_1}$  of these  $2^{k_1 T}$ -subcubes so that precisely  $R_{k_1}$  subcubes in with nonempty intersection with  $\tilde{A}_0$  remain in  $Q$ . Now consider all  $2^{-(k_1-2)T}$ -cubes intersecting the surviving set, and choose  $R_{k_1-1}$  in a similar way, etc.

If  $T$  is large enough

$$\sum_{\ell=1}^{k_1} \log_2 R_\ell = \log_2 \left| \tilde{A}_1 \right| \geq -(\alpha - \epsilon) \log_2 r \geq (\alpha - \epsilon) T (k_1 - 1)$$

the first equality being a consequence of the fact that no  $2^{-k_1 T}$ -cube can contain more than one point of  $\tilde{A}_1$ .

Set

$$M_i = \min_{i < k \leq k_1} \frac{1}{k - i} \sum_{\ell=i+1}^k \log_2 R_\ell.$$

Let  $1 \leq k_2 < k_1$  be the smallest integer for which  $M_{k_2} > (\alpha - s + \epsilon) T$  if such exists; otherwise set  $k_2 = k_1$ . Then a standard covering argument gives that there is some  $k_2 \leq k \leq k_1$  so that

$$\sum_{\ell=1}^k \log_2 R_\ell \leq k(\alpha - s + \epsilon) T$$

hence using the trivial bound  $R_\ell \leq 2^{dT}$  we get the inequality

$$k_2(\alpha - s + \epsilon)T + (k_1 - k_2)dT \geq (\alpha - \epsilon)T(k_1 - 1)$$

and

$$k_2 \leq k_1 \frac{d - \alpha + \epsilon}{d - \alpha + s - \epsilon} + O(1)$$

(explicitly, the  $O(1)$  term is  $(\alpha - \epsilon)/(d - \alpha + s - \epsilon)$ ).

Now let  $I$  be any  $2^{-k_2 T}$ -cube intersecting  $\tilde{A}_1$ , and let  $\rho_i$  be the normalized counting measure on  $\tilde{A}_1 \cap I$  as above. Then as  $M_{k_2} \geq (\alpha - s + \epsilon)T$  for any  $2^{-\ell T}$ -cube  $J \subset I$  for  $k_2 \leq \ell \leq k_1$

$$\rho_I(J) = \prod_{\ell'=k_2+1}^{\ell} R_{\ell'}^{-1} \leq 2^{-(\ell-k_2)(\alpha-s+\epsilon)T}$$

and  $\rho_I$  is a  $(C_\epsilon, \alpha - s)$ -regular measure on  $I$  at scale  $r$ , for a suitably chosen constant  $C_\epsilon$ ; note also that  $I$  is a cube of diameter  $C'_\epsilon r^\beta$  for  $\beta = k_2/k_1 = \frac{d-\alpha+\epsilon}{d-\alpha+s-\epsilon}$ .  $\square$

**Lemma 5.3.** *Let  $\rho$  be a  $(C, \alpha)$ -regular probability measure at scale  $r$  on  $B \subset \mathbb{R}^d$ . Then for any  $\epsilon > 0$  there is a  $r$ -separated subset  $A \subset \text{supp } \rho$  so that the uniform measure on  $A$  (i.e.  $\mu_A = \frac{1}{|A|} \sum_{a \in A} \delta_a$ ) is  $(C_\epsilon, \alpha - \epsilon)$ -regular at scale  $r$  on  $B$ .*

*Proof.* For simplicity of notations, we may assume without loss of generality that  $\text{diam } B = 1$ . Choose randomly a sequence  $a_1, a_2, \dots, a_L \in \text{supp } \rho$  with  $L = r^{-(\alpha-\epsilon)}$  with each  $a_i$  chosen i.i.d. according to  $\rho$ .

For any  $r \leq s < 1$  partition  $B$  into  $s^{-d}$ -cubes of size  $s$ ; the probability that any of these cubes contain at least  $N$  points of  $\{a_1, \dots, a_L\}$  for  $N > Ls^\alpha$  is at most

$$(5.1) \quad s^{-d} \left( \sum_{n \geq N} \binom{L}{n} s^{n\alpha} \right) r \leq 2s^{-d} \frac{(Ls^\alpha)^N}{N!}.$$

Choosing  $N = \max(\frac{2d}{\epsilon}, 3Ls^{\alpha-\epsilon})$ , and using the trivial inequality  $N! > (N/3)^N$  valid for all  $N > 1$ , we see that

$$(5.1) \leq 2s^{-d+\epsilon N} \leq 2s^d.$$

As  $2 \sum_{k=1}^{\infty} 10^{-kd} < 1$ , with positive probability for all  $k$ , no cube of size  $10^{-k}$  will contain more than  $\max(\frac{2d}{\epsilon}, 3L10^{-k(\alpha-\epsilon)})$  points of the set  $\{a_1, \dots, a_L\}$ . Apply this to  $k_0 = \lceil \log_{10}(1/r) \rceil$ , and obtain an  $r$  separated subset  $A \subset \{a_1, \dots, a_L\}$  of size  $|A| \geq \frac{\epsilon L}{2d}$  so that for any  $10^{-\ell}$ -cube  $Q$  with  $\ell < k_0$ ,

$$|Q \cap A| \leq 6d\epsilon^{-1} L 10^{-k(\alpha-\epsilon)} \leq 12d^2 \epsilon^{-2} |A| 10^{-k(\alpha-\epsilon)}$$

establishing the lemma.  $\square$

Closely related to the notion of  $(C, \alpha)$ -regular measures introduced in Definition 5.1 is the notion of  $\alpha$ -energy of a measure  $\rho$ , denoted by  $\mathcal{E}_\alpha(\rho)$ , which we

define for a compactly supported measure  $\rho$  on  $\mathbb{R}^d$  and  $\alpha < d$  by:

$$\mathcal{E}_\alpha(\rho) = \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \frac{d\rho(x) d\rho(y)}{|x-y|^\alpha}.$$

If  $\rho$  is  $(C, \alpha + \epsilon)$ -regular on a set  $B$  at all scales, then

$$\mathcal{E}_\alpha(\rho) = \alpha \iint \frac{\mu(\mathbf{B}_{x,r})}{r^{\alpha+1}} d\mu(x) dr \leq C(\text{diam } B)^{-\alpha-\epsilon} \alpha \epsilon^{-1}.$$

The energy  $\mathcal{E}_\alpha(\rho)$  can be given also in terms of the Fourier transform of  $\rho$ , up to an implicit constant that tends to  $\infty$  as  $\alpha \rightarrow d$  (cf. [25, 12.12]):

$$(5.2) \quad \mathcal{E}_\alpha(\rho) \asymp \int_{\mathbb{R}^d} |\hat{\rho}(\xi)|^2 (1 + |\xi|)^{\alpha-d} d\xi.$$

If  $\mathcal{E}_\alpha(\rho) < \infty$  then any set of positive  $\rho$  measure has Hausdorff dimension  $\geq \alpha$  (for this and further information about  $\alpha$ -energy, see [25]).

A simple way to adapt this notion to our ‘‘coarse’’ setup, where we do not care about the details of how  $\rho$  behaves at scales smaller than  $r$  is to smoothen it by convolving with an appropriate kernel. Let  $\Phi$  be a fixed radially symmetric nonnegative smooth function on  $\mathbb{R}^d$  with  $\|\Phi\|_1 = 1$  supported on  $\mathbf{B}_{0,1}$ , and set for  $r > 0$

$$\Phi_r(x) = r^{-d} \Phi(r^{-1}x).$$

Then instead of using the possibly atomic measure  $\rho$ , we can consider its smoothed version  $\rho' = \rho * \Phi_r$ . In particular, if  $\rho$  is  $(C, \alpha + \epsilon)$ -regular at scale  $r$  on a subset  $B \subset \mathbb{R}^d$  then

$$\mathcal{E}_\alpha(\rho * \Phi_r) \ll C(\text{diam } B)^{-\alpha-\epsilon} \alpha \epsilon^{-1}$$

with the implicit parameter depending only on  $d$  and the choice of  $\Phi$ .

## 6. STRUCTURE OF THE SET OF $t$ -LARGE FOURIER COEFFICIENTS

Fix some probability measure  $\mu_0 \in \text{Prob}(\mathbb{T}^d)$  and consider the sequence

$$\mu_n = \nu^{*n} * \mu_0$$

and the following sets of ‘‘large’’ coefficients

$$A_{t,n} = \left\{ b \in \mathbb{Z}^d : |\hat{\mu}_n(b)| > t \right\}.$$

Our goal in this section is to obtain the following result:

**Theorem 6.1.** *There exist constants  $\kappa_1 > \kappa_2 > 0$ ,  $L_2 > L_1 > 1$ ,  $p, C < \infty$  depending on  $\nu$  only, so that if for some  $t_0 \in (0, 1/2)$*

$$(6.1) \quad |\hat{\mu}_{n_0}(a_0)| \geq t_0$$

*then for any  $m > C(1 + |\log t_0|)$  one has*

$$\mathcal{N} \left( A_{t_0^p, n_0 - m_1} \cap \mathbf{B}_{0,N}; M \right) > t_0^p \cdot \left( \frac{N}{M} \right)^d$$

for some  $m_1 \leq m$ ,  $N$  in the range  $L_1^m < \frac{N}{\|a_0\|} < L_2^m$ , and  $M$  in the range  $\left(\frac{N}{\|a_0\|}\right)^{1-\kappa_1} < \frac{M}{\|a_0\|} < \left(\frac{N}{\|a_0\|}\right)^{1-\kappa_2}$ .

The proof of Theorem 6.1 involves the following steps.

**Proposition 6.2** (Initial dimension). *There exist  $\alpha_{ini}, C_1 > 0$  depending only on  $\nu$  so that for any measure  $\mu_0$  on  $\mathbb{T}^d$ , if  $\mu_n = \nu^{*n} * \mu_0$  satisfies for  $n = n_0, t_0 < 1/2$  that*

$$|\hat{\mu}_{n_0}(a_0)| > t_0 > 0$$

then for any integer  $m$  with

$$n_0 > m \geq C_1(1 + |\log(t_0)|)$$

it holds that

$$(6.2) \quad \mathcal{N}(A_{t_0/2, n_0-m} \cap \mathbf{B}_{0,N}; M) \geq (N/M)^{\alpha_{ini}}$$

for  $N = \exp(3\lambda_1 m/2) \|a_0\|, M = \|a_0\|$ .

**Proposition 6.3** (Improving the large scale dimension). *Given  $\alpha_{ini} > 0$  and  $\alpha_{high} < d$  there exists  $\alpha_{inc}, c_2, C > 0$  (depending on  $\nu$ ) so that if for some  $1/2 > t > 0, 1 \leq M < N$  with*

$$(6.3) \quad \log(N/M) > c_2 \log(1/t) \quad \text{and} \quad n \geq c_2 \log(N/M)$$

it holds that

$$\mathcal{N}(A_{t,n} \cap \mathbf{B}_{0,N}; M) > \left(\frac{N}{M}\right)^\alpha \text{ for some } \alpha_{ini} \leq \alpha \leq \alpha_{high},$$

then there are  $m, M', N'$  with

$$m \leq c_2 \log(N/M) \quad N' \leq N(N/M)^{c_2} \quad (N'/M') \geq (N/M)^{1/c_2}$$

so that

$$(6.4) \quad \mathcal{N}(A_{t,n-m} \cap \mathbf{B}_{0,N'}; M') > \left(\frac{N'}{M'}\right)^{\alpha+\alpha_{inc}}$$

for  $t' = Ct^{4d}$ .

Iterating this proposition we obtain:

**Corollary 6.4** (of Proposition 6.3). *Given  $\alpha_{ini} > 0$  and  $\alpha_{high} < d$  there exists  $c_3, C_3 > 0$  so that if for some  $1/2 > t > 0, 1 \leq M < N$  with  $\log(N/M) > c_3 \log(1/t)$  and  $n \geq c_3 \log(N/M)$  it holds that*

$$(6.5) \quad \mathcal{N}(A_{t,n} \cap \mathbf{B}_{0,N}; M) > \left(\frac{N}{M}\right)^{\alpha_{ini}},$$

then there are  $m, M', N'$  with

$$m \leq c_3 \log(N/M) \quad N' \leq N(N/M)^{c_3} \quad (N'/M') \geq (N/M)^{1/c_3}$$

so that

$$(6.6) \quad \mathcal{N}(A_{t',n-m} \cap \mathbb{B}_{0,N'}; M') > \left(\frac{N'}{M'}\right)^{\alpha_{high}}$$

for  $t' = t^{C_3}$ .

**Proposition 6.5** (High dimension to positive density at large scales). *There exist  $\alpha_{high}, c_4, \kappa_4 > 0$  depending only on  $\nu$  and  $q_4$  depending on  $d$  with the following properties. Assume that for some  $1/2 > t > 0$ ,  $1 \leq M < N$  with  $\log(N/M) > c_4 \log(1/t)$  and  $n \geq c_4 \log(N/M)$  it holds that*

$$\mathcal{N}(A_{t,n} \cap \mathbb{B}_{0,N}; M) > \left(\frac{N}{M}\right)^{\alpha_{high}}.$$

then there are  $m, M', N'$  with

$$m \leq c_4 \log(N/M) \quad N' \leq N(N/M)^{c_4} \quad (N'/M') \geq (N/M)^{1/c_4}$$

such that

$$\mathcal{N}(A_{t_1,n-m} \cap \mathbb{B}_{0,N'}; M') > c_4^{-1} t_1^{\kappa_4} \left(\frac{N'}{M'}\right)^d$$

for  $t_1 = c_4^{-1} t^{q_4}$ .

Let us deduce Theorem 6.1 from the above propositions.

*Proof.* Suppose  $|\widehat{\mu}_{n_0}(a_0)| \geq t_0$ . Then by Proposition 6.2 there are  $\alpha_{ini}, c_1$  so that

$$\mathcal{N}(A_{t_0/2, n_0-m_1} \cap \mathbb{B}_{0,N_1}; M_1) \geq (N_1/M_1)^{\alpha_{ini}}$$

for  $N_1 = \exp(3\lambda_1 m_1/2) \|a_0\|$ ,  $M_1 = \|a_0\|$  provided  $n_0 \geq m_1 \geq C_1(1 + |\log t_0|)$ .

Let  $\alpha_{high} < d$  be as in Proposition 6.5, and  $c_3, C_3$  as in Corollary 6.4, for the already chosen values of  $\alpha_{ini}, \alpha_{high}$ . Then if

$$(6.7) \quad \log(N_1/M_1) = \frac{m_1 \lambda_1}{2} > c_3(1 + |\log(t_0/2)|)$$

$$(6.8) \quad n_0 - m_1 > c_3 \log(N_1/M_1)$$

there are  $m_2 \leq c_3 \log(N_1/M_1)$  and  $N_2, M_2$  with

$$N_2 < N_1(N_1/M_1)^{c_3} \quad \frac{N_2}{M_2} \geq (N_1/M_1)^{1/c_3}$$

so that

$$\mathcal{N}(A_{t_2, n_0-m_1-m_2} \cap \mathbb{B}_{0,N_2}; M_2) > \left(\frac{N_2}{M_2}\right)^{\alpha_{high}}$$

with  $t_2 = (t_0/2)^{C_3}$ .

As long as

$$(6.9) \quad \log(N_2/M_2) > c_4(1 + \log(t_2))$$

$$(6.10) \quad n_0 - m_1 - m_2 \geq c_4 \log(N_2/M_2)$$

we can apply Proposition 6.5 and conclude that for some  $N_3, M_3$  with

$$m_3 \leq c_4 \log(N_2/M_2) \quad N_3 \leq N_2(N_2/M_2)^{c_4} \quad (N_3/M_3) \geq (N_2/M_2)^{1/c_4}$$

so that

$$\mathcal{N}(A_{t_3, n_0 - m_1 - m_2 - m_3} \cap \mathbf{B}_{0, N_3}; M_3) > c_4^{-1} t_3^{\kappa_4} \left( \frac{N_3}{M_3} \right)^d$$

with  $t_3 = (t_2)^{q_4}$ , proving the theorem.  $\square$

### 6.A. Initial Dimension and Regularity: proof of Proposition 6.2.

*Proof of Proposition 6.2.* Let  $\omega = \lambda_1/4$ , and  $C_1 > 2$  a large constant to be determined later. For any fixed  $m > C_1(1 + |\log(t_0)|)$  set

$$N = e^{(\lambda_1 + 2\omega)m} \|a_0\|, \quad R = e^{(\lambda_1 - 2\omega)m} \|a_0\|.$$

Let  $t' = \frac{t_0}{2}$  and  $n' = n_0 - m$ . Consider the following sets:

$$\begin{aligned} \mathcal{G}_{len} &= \left\{ g \in \Gamma : e^{(\lambda_1 - \omega) \cdot m} < \|g\| = \|g^{\text{tr}}\| < e^{(\lambda_1 + \omega) \cdot m} \right\}, \\ \mathcal{G}_{stat} &= \left\{ g \in \Gamma : |\hat{\mu}_{n'}(g^{\text{tr}} a_0)| \geq t' = \frac{t_0}{2} \right\}, \\ \mathcal{G}_{ang} &= \left\{ g \in \Gamma : d_{\text{ang}}(\bar{a}_0, H(g)) = d_{\text{ang}}(\theta(g), a_0^\perp) > \left( \frac{t_0}{8} \right)^{1/\tau} \right\}, \\ \mathcal{G} &= \mathcal{G}_{len} \cap \mathcal{G}_{stat} \cap \mathcal{G}_{ang}. \end{aligned}$$

By Theorem 4.3 there is  $\rho_\omega > 0$  so that (assuming  $m > m_\omega$ )

$$\nu^{*m}(\mathcal{G}_{len}) > 1 - e^{-\rho_\omega \cdot m}.$$

Our choice of  $C_1$  should guarantee  $m > m_\omega$  and

$$e^{-\rho_\omega \cdot m} < \frac{t_0}{8}.$$

There exists  $m_1$  so that for  $m > m_1$  Lemma 4.5 gives

$$\nu^{*m}(\mathcal{G}_{ang}) > 1 - \frac{t_0}{8}.$$

Finally, the fact that  $\mu_{n_0} = \nu^{*m} * \mu_{n'}$  gives

$$\nu^{*m}(\mathcal{G}_{stat}) > \frac{t_0}{2}.$$

Therefore

$$\nu^{*m}(\mathcal{G}) > \frac{t_0}{2} - \frac{t_0}{8} - \frac{t_0}{8} = \frac{t_0}{4}.$$

Since  $\|g^{\text{tr}}\| = \|g\|$ , and  $d_{\text{ang}}(\bar{x}, H(g^{\text{tr}})) = d_{\text{ang}}(\theta(g), \bar{x}^\perp)$ , by Lemma 4.1 (2) every  $g \in \mathcal{G} \subset \mathcal{G}_{ang} \cap \mathcal{G}_{len}$  has

$$\|g^{\text{tr}} a_0\| \geq \|g^{\text{tr}}\| \cdot \|a_0\| \cdot d_{\text{ang}}(\bar{a}, H(g^{\text{tr}})) > e^{(\lambda_1 - \omega)m} \|a_0\| \cdot \left( \frac{t_0}{8} \right)^{1/\tau}.$$

If  $m > (\omega \cdot \tau)^{-1} \cdot \log(8/t_0)$ , which is true for large  $C_1$ , then the the RHS above is bigger than  $R$ . It is also clear that if  $g \in \mathcal{G}_{len}$ ,  $\|g^{\text{tr}} a_0\| < N$ . So we get

$$R < g^{\text{tr}} a_0 < N \quad (g \in \mathcal{G})$$

while  $g^{\text{tr}}a_0 \in A_{t',n'}$  for any  $g \in \mathcal{G} \subset \mathcal{G}_{\text{stat}}$ . Thus for any  $g \in \mathcal{G}$  we have

$$g^{\text{tr}}a_0 \in A := A_{t',n'} \cap (\mathbb{B}_{0,N} \setminus \mathbb{B}_{0,R}) = \{R < \|b\| \leq N : |\hat{\mu}_{n'}(b)| > t'\}.$$

Let  $D$  be the projection of  $A$  to  $\mathbb{P}^{d-1}$ . Then

$$\nu^{*m} \{g : g^{\text{tr}}\bar{a}_0 \in D\} \geq \frac{t_0}{4}$$

and

$$\mathcal{N}(D; e^{-(\lambda_1 - 2\omega)m}) \leq \mathcal{N}(A; \|a_0\|).$$

It follows that

$$\begin{aligned} \nu^{*m} \{g : g^{\text{tr}}\bar{a}_0 \in D\} &\leq \mathcal{N}(D; e^{-(\lambda_1 - 2\omega)m}) \cdot \max_{\bar{y} \in \mathbb{P}^{d-1}} \nu^{*m} \{g : d_{\text{ang}}(g^{\text{tr}}\bar{a}_0, \bar{y}) < e^{-(\lambda_1 - 2\omega)m}\} \\ &\leq \mathcal{N}(D; e^{-(\lambda_1 - 2\omega)m}) e^{-\tau' m}. \end{aligned}$$

For some  $\tau' > 0$  depending only on  $\nu$ . It follows that

$$\mathcal{N}(A; \|a_0\|) \geq \frac{t_0}{4} e^{\tau' m} \geq e^{\tau' m/2}$$

if  $C_1$  is large enough. □

### 6.B. Bootstrap of large scale dimension: proof of Proposition 6.3.

A central step in the proof of Theorem 6.1 is the bootstrap procedure, which allows us to increase the large-scale “dimension” of the set of large Fourier coefficients from  $\alpha$  to  $\alpha + \alpha_{\text{inc}}$ . In order to show this we employ the following projection theorem due to the first author which implicitly can be found in [4], and are proved explicitly in [5].

**Theorem 6.6** ([5, Thm. 5]). *For any  $\alpha_0, \kappa > 0$  and  $d \geq 2$  there are  $\alpha_\Delta, \epsilon, r_0 > 0$  such that such that the following holds for  $0 < r < r_0$  and  $\alpha_0 < \alpha < d - \alpha_0$ : Let  $\eta$  be a probability measure on  $\mathbb{P}^{d-1}$  s.t.*

$$(6.11) \quad \max_{\bar{y}} \eta(V(y^\perp, \rho)) < \rho^\kappa \quad \text{if } r < \rho < r^{\tau_0}$$

Let  $E \subset [0, 1]^d$  be a  $r$ -separated set with  $|E| > r^{-\alpha}$  and a non-concentration property

$$\max_x |E \cap \mathbb{B}_{x,\rho}| < \rho^\kappa |E| \quad \text{if } r < \rho < r^{\tau_0}.$$

Then there exist  $D \subset \mathbb{P}^{d-1}$  and  $E' \subset E$  with

$$\eta(D) > 1 - r^\epsilon \quad |E'| > r^\epsilon |E|$$

So that

$$\mathcal{N}(\pi_\theta(E''); r) > r^{-(\alpha + \alpha_\Delta)/d}$$

whenever  $\theta \in D$  and  $E'' \subset E'$  satisfies  $|E''| > r^{2\epsilon} |E|$ .

**Lemma 6.7.** *For any  $\epsilon > 0$ , there is a  $C_\epsilon > 0$  so that the following holds. Let  $\mu$  be a measure on  $\mathbb{T}^d$  and*

$$A_t(\mu) = \left\{ b \in \mathbb{Z}^d : |\widehat{\mu}(b)| > t \right\}.$$

Assume that for some  $N > M, \alpha$

$$\mathcal{N}(A_t(\mu) \cap \mathbf{B}_{0,N}; M) \geq \left( \frac{N}{M} \right)^\alpha.$$

Then there is a  $M < N' < N$  with

$$\log \frac{N'}{M} > \left( \frac{d - \alpha + \epsilon}{d - \alpha + 8\epsilon} \right) \log \frac{N}{M}$$

so that  $A_{t/4}(\mu) \cap \mathbf{B}_{0,N'}$  contains a subset which is  $(C_\epsilon t^{-2}, \alpha - 10\epsilon)$ -regular at scale  $M$ .

*Proof.* By Lemma 5.2, there is a point  $x \in \mathbf{B}_{0,N}$  so that  $A \cap \mathbf{B}_{x,N'}$  supports a probability measure  $\rho$  which is  $(C_\epsilon, \alpha - 9\epsilon)$ -regular measure at scale  $M$  with

$$\frac{N'}{M} = \left( \frac{N}{M} \right)^{\frac{d - \alpha + \epsilon}{d - \alpha + 8\epsilon}}.$$

Replacing  $C_\epsilon$  by  $4C_\epsilon$  we may assume all  $b \in \text{supp } \rho$  satisfies that  $\widehat{\mu}(b)$  lie in a single quadrant of  $\mathbb{C}$ , and hence

$$\left| \sum_b \rho(b) \widehat{\mu}(b) \right| \geq \frac{t}{\sqrt{2}}.$$

By Cauchy-Schwartz:

$$\begin{aligned} \sum_{b, b_1} \widehat{\mu}(b - b_1) \rho(b) \rho(b_1) &= \int_{\mathbb{T}^d} \left| \sum_b e(b \cdot x) \rho(b) \right|^2 dx \\ &\geq \left| \int_{\mathbb{T}^d} \sum_b e(b \cdot x) \rho(b) dx \right|^2 \\ &= \left| \sum_b \widehat{\mu}(b) \rho(b) \right|^2 \geq \frac{t^2}{2} \end{aligned}$$

hence

$$(6.12) \quad \rho * \rho(A_{t/4}(\mu)) \geq \frac{t^2}{4}.$$

Let  $\rho_2$  be the probability measure  $\rho * \rho|_{A_{t/4}(\mu)}$ . As  $\rho$  was  $(4C_\epsilon, \alpha - 9\epsilon)$ -regular on  $\mathbf{B}_{x,N'}$ , the measure  $\rho * \rho$  is  $(2^{d+2}C_\epsilon, \alpha - 9\epsilon)$ -regular on  $\mathbf{B}_{0,2N'}$ , hence by (6.12)

$$\rho_2 \text{ is } (2^{d+4}C_\epsilon t^{-2}, \alpha - 9\epsilon)\text{-regular on } B(0, 2N').$$

By Lemma 5.3, there is some

$$\tilde{A} \subset \text{supp } \rho_2 \subset A_{t/4}(\mu) \cap \mathbf{B}_{0,N'}$$



which is  $M$ -separated and  $(C'_\epsilon t^{-2}, \alpha - 10\epsilon)$ -regular on  $\mathbf{B}_{0, N'}$ .  $\square$

**Lemma 6.8.** *Given  $\alpha_{ini} > 0$  and  $\alpha_{high} < d$  there exists  $\alpha_{inc}, c_6, C > 0$  (depending on  $\nu$ ) so that if for some  $1/2 > t > 0$ ,  $1 \leq M < N$  with*

$$(6.13) \quad \log(N/M) > c_6 \log(1/t) \quad \text{and} \quad n \geq c_6 \log(N/M)$$

it holds that

$$\mathcal{N}(A_{t,n} \cap \mathbf{B}_{0,N}; M) > \left(\frac{N}{M}\right)^\alpha \text{ for some } \alpha_{ini} \leq \alpha \leq \alpha_{high},$$

then there are  $m, M', N'$  with  $M' \geq M$

$$m \leq c_6 \log(N/M) \quad N' \leq N(N/M)^{c_6} \quad (N'/M') \geq (N/M)^{1/c_6}$$

and  $\xi \in \mathbb{P}^{d-1}$  so that if  $R$  denotes the ‘‘rectangle’’  $\mathbf{B}_{0, N'} \cap \mathbf{Nbd}_{M'}(\xi)$

$$(6.14) \quad \mathcal{N}(A_{t', n-m} \cap R; M') > \left(\frac{N'}{M'}\right)^{(\alpha + 2\alpha_{inc})/d}$$

for  $t' = Ct^4$ .

*Proof.* Let  $\alpha_\Delta$  be as in Theorem 6.6 for  $\alpha_0 = \kappa = \min(\alpha_{ini}, d - \alpha_{high})/2$ . By Lemma 6.7 applied with  $\epsilon = \alpha_\Delta/20$ , there is an  $M < N_1 < N$  with

$$\log(N_1/M) > c \log(N/M)$$

and an  $M$ -separated subset

$$A \subset A_{t^2/4, n} \cap \mathbf{B}_{0, N_1}$$

which is  $(Ct^{-2}, \alpha - \alpha_\Delta/2)$ -regular at scale  $M$  on  $\mathbf{B}_{0, N_1}$ ; in particular

$$|A| > C^{-1} t^2 \left(\frac{N_1}{M}\right)^{\alpha - \alpha_\Delta/2}.$$

Both the constants  $c$  and  $C$  depend only on  $\alpha_{ini}$  and  $\alpha_{high}$  (and  $\alpha_\Delta$  which is determined by these two quantities).

Let  $\omega > 0$  be small (specifically, we require that  $\omega < \min(\lambda_1 - \lambda_2, \lambda_1, \alpha_\Delta)/20$ ) and let  $m$  be the smallest integer so that

$$e^{(\lambda_1 - \lambda_2 - 2\omega) \cdot m} > \frac{N_1}{M}.$$

Let  $n' = n - m$  and set  $M', N'$  by

$$(m1) \quad N' = e^{(\lambda_1 + \omega) \cdot m} \cdot N_1,$$

$$(m2) \quad M' = e^{(\lambda_1 - \omega) \cdot m} \cdot M,$$

then also

$$(m3) \quad M' \leq e^{(\lambda_2 + \omega) \cdot m} \cdot N_1 \leq M'.$$

assuming the constant  $c_6$  in (6.13) is sufficiently large, we will have that  $m$  is greater or equal to the constant  $m_0(\omega)$  in Theorem 4.3. Invoking that theorem, we conclude that the set

$$\mathcal{G}_{len} = \left\{ g \in \Gamma : \left| \lambda_i - \frac{1}{m} \log \sigma_i(g) \right| < \omega \text{ for } i = 1, 2 \right\}.$$

satisfies

$$(6.15) \quad \nu^{*m}(\mathcal{G}_{len}) > 1 - e^{-\rho_\omega \cdot m}$$

Conditions (m1)–(m3) imply that for any  $g \in \mathcal{G}_{len}$

$$g^{\text{tr}}(\mathbf{B}_{0,N_1}) \subset \mathbf{B}_{0,N'} \cap \text{Nbd}_{M'}(\xi) \quad \text{with} \quad \xi = \theta(g^{\text{tr}}),$$

i.e. the linear transformation  $g^{\text{tr}}$  maps the ball  $\mathbf{B}_{0,N_1}$  into a cylinder of length  $2N'$  and base of radius  $M'$ .

Let  $\eta \in \text{Prob}(\mathbb{P}^{d-1})$  denote the distribution of  $\theta(g)$  where  $g \in \Gamma$  is distributed according to  $\nu^{*m}$ , i.e.,  $\eta(\Theta) = \nu^{*m} \{g \in \Gamma : \theta(g) \in \Theta\}$ . Lemma 4.5 provides the regularity of  $\eta$  as in condition (6.11) of Theorem 6.6.

Let  $E = N_1^{-1} \cdot A \subset \mathbf{B}_{0,1} \subset \mathbb{R}^d$  and  $r = M/N_1$ . Theorem 6.6, applied for  $r = M/N'$  and  $\epsilon = \alpha_\Delta/10$ , gives us a set  $E' \subset E$  with  $|E'| > r^{\alpha_\Delta/10} |E|$  and  $\Theta \subset \mathbb{P}^{d-1}$  with  $\eta(\Theta) > 1 - r^{\alpha_\Delta/10}$  so that

$$(6.16) \quad \mathcal{N}(\pi_\theta(E''); r) \geq r^{-(\alpha + \frac{1}{2}\alpha_\Delta)/d} \quad \forall E'' \subset E', \theta \in \Theta \text{ with } |E''| > r^{\alpha_\Delta/10} |E'|.$$

Let  $B = N_1 E'$  and

$$\mathcal{G}_{proj} = \{g \in \Gamma : \theta(g) \in \Theta\}.$$

We have

$$(6.17) \quad \nu^{*m}(\mathcal{G}_{proj}) = \eta(\Theta) > 1 - r^{\alpha_\Delta/10}.$$

Since  $b \in B \subset A \subset A_{t^2/4,n}$ , we have that  $|\hat{\mu}_n(b)| > \frac{1}{4}t^2$  for all  $b \in B$ ; by reducing  $B$  slightly we may also assume that  $|B|^{-1} |\sum_{b \in B} \hat{\mu}_n(b)| \geq \frac{1}{8}t^2$ . Using the identity  $\mu_n = \nu^{*m} * \mu_{n'}$  (recall that  $n' = n - m$ ) and the Cauchy-Schwartz inequality we may conclude that

$$\begin{aligned} \sum_{g \in \Gamma} \nu^{*m}(g) \cdot \frac{1}{|B|} \sum_{b \in B} |\hat{\mu}_{n'}(g^{\text{tr}}b)|^2 &\geq \left| \frac{1}{|B|} \sum_{g \in \Gamma} \sum_{b \in B} \nu^{*m}(g) \hat{\mu}_{n'}(g^{\text{tr}}b) \right|^2 \\ &= \left| \frac{1}{|B|} \sum_{b \in B} \hat{\mu}_n(b) \right| > 2^{-6}t^4 \end{aligned}$$

and therefore the set

$$\mathcal{G}_{stat} = \left\{ g \in \Gamma : \frac{1}{|B|} \sum_{b \in B} |\hat{\mu}_{n'}(g^{\text{tr}}b)|^2 > 2^{-7}t^4 \right\}$$

has

$$(6.18) \quad \nu^{*m}(\mathcal{G}_{stat}) > 2^{-7}t^4.$$

Note that for each  $g \in \mathcal{G}_{stat}$  the set

$$(6.19) \quad B_g = \{b \in B : |\hat{\mu}_{n'}(g^{\text{tr}}b)|^2 > 2^{-8}t^4\}$$

has  $|B_g| > 2^{-8}t^4 \cdot |B|$ . Let  $\mathcal{G} = \mathcal{G}_{len} \cap \mathcal{G}_{proj} \cap \mathcal{G}_{stat}$ . From (6.18), (6.17) and (6.15) we have

$$\nu^{*m}(\mathcal{G}) > 2^{-7}t^4 - r^{\alpha\Delta/10} - e^{-\rho\omega \cdot m} > 2^{-8}t^4, \quad \text{where} \quad r = \frac{M}{N_1}.$$

assuming  $r^{\alpha\Delta/10}, e^{-\rho\omega \cdot m} < 2^{-9}t^4$  which is guaranteed by taking  $c_6$  large enough.

Moreover, for any  $g \in \mathcal{G}$  we have that  $|B_g| > 2^{-8}t^4 \cdot |B|$  and (assuming as we may that  $2^{-8}t^4 > r^{\alpha\Delta/10}$ ) by (6.16)

$$\mathcal{N}(\pi_\xi(B_g); M) \geq r^{-(\alpha + \frac{1}{2}\alpha\Delta)/d} \quad \text{with } \xi = \theta(g^{\text{tr}});$$

note that by definition of  $B_g$ ,

$$(6.20) \quad g^{\text{tr}}(B_g) \subset A_{\frac{t_1^2}{16}, n'}.$$

Since also  $g \in \mathcal{G}_{len}$ ,  $g^{\text{tr}}(B_g) \subset \mathbf{B}_{0, N'} \cap \mathbf{Nbd}_{M'}(\xi)$  and

$$\mathcal{N}(g^{\text{tr}}(B_g); M') \geq \mathcal{N}(\pi_\xi(B_g); M),$$

which in view of (6.20) implies (6.14).  $\square$

**Lemma 6.9.** *Let  $1/2 > t_1 > 0, M_1 < N_1$  and  $n_1$  satisfy*

$$(6.21) \quad n_1, \log(N/M) > c_7 \log(1/t_1)$$

with  $c_7$  depending on  $\nu$ . Let  $\xi \in \mathbb{P}^{d-1}$  and let  $R$  be the ‘‘rectangle’’  $R = \mathbf{B}_{0, N_1} \cap \mathbf{Nbd}_{M_1}(\xi)$ . Then there are  $m_2, M_2, N_2$  with

$$m_2, |\log(N_1) - \log(N_2)|, |\log(M_1) - \log(M_2)| \leq c_7 \log(1/t_1)$$

so that for  $t_2 = (t_1/8)^{2d}$

$$(6.22) \quad \mathcal{N}(A_{t_2, n_1 - m_2} \cap \mathbf{B}_{0, N_2}; M_2) \geq c_7^{-1} t_1^{\kappa_7} \mathcal{N}(A_{t_1, n_1} \cap R; M_1)^d$$

where  $\kappa_7$  also depends only on  $\nu$ .

*Proof.* Let  $\omega = (\lambda_1 - \lambda_2)/10$ , and let  $m_2$  be such that the sets

$$\begin{aligned} \mathcal{G}_{len} &= \left\{ \vec{g} \in \Gamma^d : \left| \lambda_i - \frac{1}{m_2} \log \sigma_i(g_j) \right| < \omega \text{ for } i = 1, 2 \text{ and } j = 1, \dots, d \right\} \\ \mathcal{G}_{ang} &= \left\{ \vec{g} \in \Gamma^d : d_{\text{ang}}(\xi, H(g_j^{\text{tr}})) \geq 2e^{-\omega m_2} \text{ for } j = 1, \dots, d \right\} \\ \mathcal{G}_{vol} &= \left\{ \vec{g} \in \Gamma^d : \text{vol}(\theta(g_1^{\text{tr}}), \dots, \theta(g_d^{\text{tr}})) > \left( \frac{t_1}{8} \right)^{2dp} \right\} \quad (p \text{ as in Lemma 4.6}) \end{aligned}$$

satisfy

$$(6.23) \quad \min((\nu^{*m_2})^d(\mathcal{G}_{len}), (\nu^{*m_2})^d(\mathcal{G}_{ang}), (\nu^{*m_2})^d(\mathcal{G}_{vol})) \geq 1 - \frac{1}{6}(t_1/4)^{2d}.$$

By Lemma 4.6 and Theorem 4.3 and Theorem 4.4 one can find such  $m_2$  with  $m_2 < \tau_1 |\log t_1|$  for some constant  $\tau_1$ . In particular if the constant  $c_7$  of (6.21) is sufficiently large,  $n_2 = n_1 - m_2 > 0$ , which we shall assume henceforth.

Let  $E \subset A_{t_1, n_1} \cap R$  be such that

(E1) for every  $b, b' \in E$

$$d_{\text{ang}} \left( \xi, \frac{b - b'}{|b - b'|} \right) \leq e^{-\omega m_2}$$

(E2)  $|\sum_{b \in E} \widehat{\mu}_{n_1}(b)| \geq \frac{1}{2} t_1 |E|$ .

Clearly, by (6.28) one can find such  $E$  so that

$$|E| \geq c e^{-d\omega m_2} \mathcal{N}(A_{t_1, n_1} \cap R; M_1)$$

where  $c$  is some constant depending only on  $d$ . Note that in order to satisfy (E2) one can e.g. take  $E$  so that  $\{\widehat{\mu}_{n_1}(b) : b \in E\}$  lie in a single quadrant of  $\mathbb{C}$ .

Write  $G(x) = |E|^{-1} \sum_g \nu^{*m_2}(g) \sum_{b \in E} e_{g^{\text{tr}b}}(x)$ . Then

$$\begin{aligned} 2^{-2d} t_1^{2d} &\leq \left| \frac{1}{|E|} \sum_{b \in E} \widehat{\mu}_{n_1}(b) \right|^{2d} = \left| \frac{1}{|E|} \sum_g \sum_{b \in E} \nu^{*m_2}(g) \widehat{\mu}_{n_2}(g^{\text{tr}b}) \right|^{2d} \\ &= \left| \int G(x) d\mu_{n_2}(x) \right|^{2d} \leq \int |G(x)|^{2d} d\mu_{n_2}(x) \\ (6.24) \quad &= \frac{1}{|E|^{2d}} \sum_{g_1, \dots, g_{2d}} \nu^{*m_2}(g_1) \dots \nu^{*m_2}(g_{2d}) \cdot \\ &\quad \cdot \sum_{b_1, \dots, b_{2d} \in E} \widehat{\mu}_{n_2}(g_1^{\text{tr}b_1} + \dots + g_d^{\text{tr}b_d} \\ &\quad \quad \quad - g_{d+1}^{\text{tr}b_{d+1}} - \dots - g_{2d}^{\text{tr}b_{2d}}). \end{aligned}$$

Set  $\Sigma_{(g_1, \dots, g_d)}(\vec{b}) = \sum_{i=1}^d g_i^{\text{tr}b_i}$ . Fix  $(g_{d+1}, \dots, g_{2d}) \in \mathcal{G}_{\text{len}}$  and  $b_{d+1}, \dots, b_{2d} \in E$  with

(6.25)

$$|E|^{-d} \left| \sum_{g_1, \dots, g_d} \nu^{*m_2}(g_1) \dots \nu^{*m_2}(g_d) \sum_{\vec{b} \in E^d} \widehat{\mu}_{n_2}(\Sigma_{(g_1, \dots, g_d)}(\vec{b}) - \mathbf{b}) \right| \geq \left( \frac{t_1}{4} \right)^{2d}$$

where  $\mathbf{b} = g_{d+1}^{\text{tr}b_{d+1}} + \dots + g_{2d}^{\text{tr}b_{2d}}$ . Such a choice exists in view of the estimate (6.23) on the measure of  $\mathcal{G}_{\text{len}}$  and (6.24).

Set

$$\mathcal{G}_{\text{stat}} = \left\{ \vec{g} \in \Gamma^d : |E|^{-d} \left| \sum_{\vec{b} \in E^d} \widehat{\mu}_{n_2}(\Sigma_{\vec{g}}(\vec{b}) - \mathbf{b}) \right| > \frac{1}{2} \left( \frac{t_1}{4} \right)^{2d} \right\}.$$

In view of (6.25),  $(\nu^{*m_2})^d(\mathcal{G}_{\text{stat}}) \geq \frac{1}{2} (t_1/4)^{2d}$ , hence by (6.23) the set  $\mathcal{G} = \mathcal{G}_{\text{stat}} \cap \mathcal{G}_{\text{vol}} \cap \mathcal{G}_{\text{len}} \cap \mathcal{G}_{\text{ang}}$  is nonempty. Let  $t_2 = (t_1/8)^{2d}$ .

We claim that if

$$t_2^p e^{(\lambda_1 - 2\omega)m_2} > 4d e^{(\lambda_2 + \omega)m_2}$$

then for any  $\vec{g} \in \mathcal{G}$ ,

$$(6.26) \quad \mathcal{N}(A_{t_2, n_2} \cap (\Sigma_{\vec{g}}(E^d) - \mathbf{b}); M_2) \geq t_2 |E|^d$$

with  $M_2 = \frac{1}{4}t_2^p e^{(\lambda_1 - 2\omega)m_2} M_1$ ; note that  $\vec{g}, (g_{d+1}, \dots, g_{2d}) \in \mathcal{G}_{len}$  and  $E \subset \mathbf{B}_{0, N_1}$  imply

$$\Sigma_{\vec{g}}(E^d) - \mathbf{b} \subset \mathbf{B}_{0, N_2} \quad \text{where } N_2 = 2\sqrt{d}e^{(\lambda_1 + \omega)m_2} N_1.$$

As

$$(6.27) \quad |E| \geq ce^{-d\omega m_2} \mathcal{N}(A_{t', n_1} \cap R; M_1)$$

It follows from (6.26) and (6.27) that

$$\mathcal{N}(A_{t_2, n_2} \cap \mathbf{B}_{0, N_2}; M_2) \geq c't_2 e^{-d^2\omega m_2} \mathcal{N}(A_{t', n_1} \cap R; M_1)^d,$$

establishing Lemma 6.9 assuming the claim (6.26).

We now turn to proving (6.26). Let  $\xi_i = \theta(g_i^{\text{tr}})$  for  $i = 1, \dots, d$ . We shall use the following auxiliary expression, which is meant to approximate  $\Sigma_{\vec{g}}(\vec{b})$ :

$$\Sigma_{\vec{g}}^*(\vec{b}) = \sum_{i=1}^d \pi_{\xi_i}(g_i^{\text{tr}} b_i),$$

where we consider  $\pi_{\xi_i}$  as a rank one map  $\mathbb{R}^d \rightarrow \mathbb{R}^d$  whose image is in the vector space spanned by  $\xi_i$ . Indeed, for  $\vec{g} \in \mathcal{G}_{len}$

$$\left\| \Sigma_{\vec{g}}(\vec{b}) - \Sigma_{\vec{g}}^*(\vec{b}) \right\| \leq e^{(\lambda_2 + \omega)m_2} \sum_i \|b_i\|.$$

Let  $\vec{b}^{(i)} = (b_1^{(i)}, \dots, b_d^{(i)})$  ( $i = 1, 2$ ) be two distinct points in  $E^d$ ; assume they differ in the  $j$ th coordinate  $b_j^{(i)}$ . Write  $b_j = b_j^{(1)} - b_j^{(2)}$  as  $b'_j + b''_j$  with  $b''_j \in H(g_j)$  and  $b'_j \perp b''_j$ . As  $E$  is  $M_1$ -separated,  $\|b_j\| \geq M_1$ . Then  $d_{\text{ang}}(\xi, b_j / \|b_j\|) < e^{-\omega m_2}$  (cf. (E1)), hence as  $\vec{g} \in \mathcal{G}_{ang}$

$$d_{\text{ang}}\left(\frac{b_j}{\|b_j\|}, H(g_j)\right) \geq d_{\text{ang}}(\xi, H(g_j^{\text{tr}})) - d_{\text{ang}}\left(\frac{b_j}{\|b_j\|}, \xi\right) \geq e^{-\omega m_2}$$

and

$$|b'_j| \geq e^{-\omega m_2} |b_j| \geq \frac{1}{2}e^{-\omega m_2} M_1.$$

In these notations,  $\pi_{\xi_j}(g_j^{\text{tr}} b_j) = g_j^{\text{tr}} b'_j$ , and it follows that

$$\left\| \pi_{\xi_j}(g_j^{\text{tr}} b_j) \right\| \geq \frac{1}{2}e^{(\lambda_1 - 2\omega)m_2} \|b_j\|.$$

Then as  $\vec{g} \in \mathcal{G}_{vol}$

$$\begin{aligned} \left\| \Sigma_{\vec{g}}^*(\vec{b}^{(1)} - \vec{b}^{(2)}) \right\| &\geq \text{vol}(\Sigma_{\vec{g}}(\vec{b}^{(1)} - \vec{b}^{(2)}), \xi_1, \dots, \xi_{j-1}, \xi_{j+1}, \dots, \xi_d) \\ &= \left\| \pi_{\xi_j}(g_j^{\text{tr}} b_j) \right\| \text{vol}(\xi_1, \dots, \xi_d) \\ &\geq \frac{1}{2}t_2^p e^{(\lambda_1 - 2\omega)m_2} \|b_j\|. \end{aligned}$$

Hence

$$\left\| \Sigma_{\vec{g}}(\vec{b}^{(1)} - \vec{b}^{(2)}) \right\| \geq \frac{1}{2}t_2^p e^{(\lambda_1 - 2\omega)m_2} \|b_j\| - de^{(\lambda_2 + \omega)m_2} \|b_j\|$$

under the assumption this last expression is

$$\geq \frac{1}{4}t_2^p e^{(\lambda_1 - 2\omega)m_2} \|b_j\| \geq M_2$$

for which it follows that  $\Sigma_{\bar{g}}(E^d)$  is  $M_2$ -separated and (6.26) is proved, concluding the proof of Lemma 6.9.  $\square$

*Proof of Proposition 6.3.* Apply Lemma 6.8 to find  $N_1, M_1, m_1, n_1 = n - m_1$  with

$$m_1 \leq c_6 \log(N/M) \quad N_1 \leq N(N/M)^{c_6} \quad (N_1/M_1) \geq (N/M)^{1/c_6}$$

and a  $\xi \in \mathbb{P}^{d-1}$  so that

$$(6.28) \quad \mathcal{N}(A_{t_1, n_1} \cap R; M_1) > \left(\frac{N_1}{M_1}\right)^{(\alpha+2\alpha_{inc})/d} \quad (t_1 = Ct^4)$$

with  $R = \mathbf{B}_{0, N_1} \cap \mathbf{Nbd}_{M_1}(\xi)$ . Now apply Lemma 6.9 to find  $m_2, n_2 = n_1 - m_2, M_2, N_2$  with

$$m_2, |\log(N_1) - \log(N_2)|, |\log(M_1) - \log(M_2)| \leq c_7 \log(1/t_1)$$

so that

$$\begin{aligned} \mathcal{N}(A_{t_2, n_2} \cap \mathbf{B}_{0, N_2}; M_2) &\geq c_7 t_1^{\kappa_7} \mathcal{N}(A_{t_1, n_1} \cap R; M_1) \\ &> c_7 t_1^{\kappa_7} \left(\frac{N_1}{M_1}\right)^{\alpha+2\alpha_{inc}} \end{aligned}$$

with  $t_2 = (t_1/8)^{2d}$ . Note that by choosing  $c_2$  of (6.3) to be large enough guarantees that (6.21) holds. Moreover, if this constant  $c_2$  is large enough,

$$c_7 t_1^{\kappa_7} \left(\frac{N_1}{M_1}\right)^{\alpha+2\alpha_{inc}} > \left(\frac{N_2}{M_2}\right)^{\alpha+\alpha_{inc}}$$

establishing Proposition 6.3.  $\square$

### 6.C. From high dimension to positive density: proof of Proposition 6.5.

Underlying (and motivating) the proof of Proposition 6.5 is the following theorem of Falconer [13] regarding projection of sets. Falconer shows that if  $\eta$  is a measure on the set of directions with dimension  $\beta > 0$  then if the dimension of  $\rho$  is larger than  $d - \beta$  one has that for  $\eta$  almost every direction  $\theta$  the projection  $\rho_\theta$  of  $\rho$  in the direction  $\theta$  is absolutely continuous with respect to Lebesgue measure; we follow the treatment of this result by Peres and Schlag in [27, Sec. 6]. In fact, the argument gives a much more quantitative result connecting the  $\alpha$ -energy of  $\rho$  to the projections of  $\rho$ .

We need a version of this theorem for measures  $\rho$  which are  $(C, \alpha)$ -regular at some scale  $r$ , but are possibly singular at finer scales (indeed the measure we shall consider will be purely atomic). As we have already remarked in Section 5, this can be achieved by applying Falconer's theorem to  $\rho$  convolved with an appropriate smoothing function.

Let  $\Phi$  be a fixed radially symmetric nonnegative smooth function on  $\mathbb{R}^d$  with  $\|\Phi\|_1 = 1$  supported on  $\mathbf{B}_{0,1}$ , and set for  $r > 0$

$$(6.29) \quad \Phi_r(x) = r^{-d} \Phi(r^{-1}x).$$

Let  $\Psi : \mathbb{R} \rightarrow \mathbb{R}^+$  be the smooth compactly supported function

$$\Psi(x_1) = \int dx_2 \dots \int dx_d \Phi(x_1, x_2, \dots, x_d),$$

and define  $\Psi_r$  analogously to (6.29).

**Lemma 6.10.** *Let  $\rho$  be a probability measure on  $\mathbb{R}$ , and  $\phi$  the Radon-Nykodim derivative  $\phi = \frac{d(\rho * \Psi_r)}{dx}$ . Then for every  $r < r_1 < 1$*

$$(6.30) \quad \mathcal{N}(\text{supp } \rho; r_1) \geq (4r_1 \|\phi\|_2^2)^{-1}.$$

Moreover, for any subset  $X \subset \text{supp } \rho$ ,

$$(6.31) \quad \mathcal{N}(X; r_1) \geq \frac{\rho(X)^2}{4r_1 \|\phi\|_2^2}.$$

*Proof.* Let  $B = \text{supp } \rho + [-r, r]$ , and  $1_B$  the corresponding indicator function. Then the Lebesgue measure of  $B$  satisfies  $\lambda(B) \leq 4r_1 \mathcal{N}(\text{supp } \rho; r_1)$ . By Cauchy-Schwartz

$$1 = \int 1_B(x) \phi(x) dx \leq \|1_B\|_2 \|\phi\|_2.$$

Since  $\|1_B\|_2 = \sqrt{\lambda(B)}$  equation (6.30) follows.

To see (6.31), apply (6.30) on the probability measure  $\rho|_X$  defined by  $\rho|_X(Y) = \frac{1}{\rho(X)} \rho(X \cap Y)$ ; one has

$$\frac{d\rho|_X}{dx}(y) = \begin{cases} \frac{1}{\rho(X)} \frac{d\rho}{dx}(y) & \text{if } y \in X \\ 0 & \text{if } y \notin X \end{cases}$$

hence  $\|d\rho|_X/dx\|_2^2 \leq \rho(X)^{-2} \|d\rho/dx\|_2^2$ .  $\square$

**Proposition 6.11.** *Let  $\rho$  be a probability measure supported on the unit ball  $\mathbf{B}_{0,1}$  of  $\mathbb{R}^d$  so that  $\mathcal{E}_\alpha(\rho) < \infty$  for some  $0 < \alpha < d$ ,  $0 < r < 1$ , and let  $\eta$  be a measure on  $S^{d-1}$  such that for some  $c_\eta, \beta > 0$*

$$(6.32) \quad \eta(\mathbf{B}_{\theta, \epsilon}) \leq c_\eta \epsilon^\beta \quad \text{for every } \epsilon > r \text{ and } \theta \in S^{d-1}.$$

Then for any  $\beta' < \beta$

$$(6.33) \quad \int_\theta \int_t |\hat{\rho}_\theta(t)|^2 \left| \widehat{\Psi}_r(t) \right|^2 (1 + |t|)^{\beta' + \alpha - d} dt d\eta(\theta) \leq c_\eta C_d \int_{\mathbb{R}^d} |\hat{\rho}(x)|^2 \left| \widehat{\Phi}_r(x) \right|^2 (1 + |x|)^{\alpha - d} dx + c_\eta C(\alpha, \beta, \beta', d).$$

**Interpretation:** if  $\alpha + \beta' > d$  and  $\eta$  is  $(C, \alpha')$ -regular at scale  $r$  for  $\alpha' > \alpha$  then by (5.2) the right hand side of (6.32) is bounded from above by a constant (depending on  $\alpha, \alpha', \beta, \beta', C, \dots$ ) while the left-hand side dominates

$$\int_\theta \left\| \frac{d(\rho * \Psi_r)}{dx} \right\|_2^2 d\eta(\theta).$$

In view of Lemma 6.10, this in particular implies that for  $\eta$ -many choices of  $\theta$ , the covering number of  $\text{supp}(\rho_\theta)$  by  $r$ -intervals is large.

*Proof of Proposition 6.11.* Our proof follows closely that of [27, Prop. 6.1]. Let  $\chi$  be a smooth, compactly supported function on  $\mathbb{R}^d$  with nonnegative Fourier transform and  $\chi \equiv 1$  on  $\mathbf{B}_{0,1}$ . Then  $\rho = \rho \cdot \chi$  and hence  $\widehat{\rho} = \widehat{\rho} * \widehat{\chi}$ . It follows that  $|\widehat{\rho}|^2 < C |\widehat{\rho}|^2 * \widehat{\chi}$ ; also since  $\chi$  is smooth, compactly supported,

$$|\widehat{\chi}(\xi)| < C_N (1 + |\xi|)^{-N} \quad \text{for every } N;$$

we shall assume below that  $N \geq 2d$ . Thus

$$(6.34) \quad \begin{aligned} & \int_{S^{d-1}} \int_{\mathbb{R}} \left| \widehat{\rho}_\theta(t) \widehat{\Psi}_r(t) \right|^2 (1 + |t|)^{\beta' + \alpha - d} dt d\eta(\theta) \leq \\ & \leq C \int_{S^{d-1}} \int_{\mathbb{R}} \int_{\mathbb{R}^d} \widehat{\chi}(\theta t - x) \left| \widehat{\rho}(x) \widehat{\Phi}_r(x) \right|^2 (1 + |t|)^{\beta' + \alpha - d} dt d\eta(\theta) dx \\ & \leq C'_N \int_{\mathbb{R}^d} \left| \widehat{\rho}(t) \widehat{\Phi}_r(t) \right|^2 \int_{S^{d-1}} \int_{\mathbb{R}} (1 + |\theta t - x|)^{-N} (1 + |t|)^{\beta' + \alpha - d} dt d\eta(\theta) dx. \end{aligned}$$

We estimate the innermost integral in the last line of the above equation as follows:

$$\begin{aligned} & \int_{\mathbb{R}} (1 + |\theta t - x|)^{-N} (1 + |t|)^{\beta' + \alpha - d} dt \\ & \leq 2^d (1 + |x|)^{\beta' + \alpha - d} \int_{|x|/2 < |t| < 2|x|} (1 + |\theta t - x|)^{-N} dt + \\ & \quad + C(N, \beta', \alpha) (1 + |x|)^{-N} \\ & \leq C_{d,N} (1 + |x|)^{\beta' + \alpha - d} \left( 1 + |x| d_{\text{ang}} \left( \theta, \frac{x}{|x|} \right) \right)^{-N+d} + \\ & \quad + C(N, \beta', \alpha) (1 + |x|)^{-N}. \end{aligned}$$

Using (6.32), we have (recall that  $N > 2d$ )

$$\begin{aligned} & \int_{S^{d-1}} \left( 1 + |x| d_{\text{ang}} \left( \theta, \frac{x}{|x|} \right) \right)^{-N+d} d\eta(\theta) \leq \eta \left\{ \theta : d_{\text{ang}} \left( \theta, \frac{x}{|x|} \right) < |x|^{-1} \right\} + \\ & \quad + \sum_{k \geq 0} 2^{-(N-d)k} \eta \left\{ \theta : 2^k |x|^{-1} \leq d_{\text{ang}} \left( \theta, \frac{x}{|x|} \right) < 2^{k+1} |x|^{-1} \right\} \\ & \leq 10c_\eta \max(r, (1 + |x|)^{-1})^\beta. \end{aligned}$$

It follows that the integral on the last line of (6.34) is at most

$$(6.35) \quad \begin{aligned} & 10 C_{d,N} c_\eta \int_{|x| < r^{-\beta/\beta'}} \max(r, (1 + |x|)^{-1})^\beta \left| \widehat{\rho}(x) \widehat{\Phi}_r(x) \right|^2 (1 + |x|)^{\beta' + \alpha - d} d\eta(x) \\ & \quad + 10 C_{d,N} c_\eta \int_{|x| > r^{-\beta/\beta'}} r^\beta \left| \widehat{\Phi}_r(x) \right|^2 (1 + |x|)^{\beta' + \alpha - d} d\eta(x) \\ & \quad + C'(N, \beta', \alpha). \end{aligned}$$



For  $|x| < r^{-\beta/\beta'}$  one has the trivial inequality

$$\max(r, (1 + |x|)^{-1})^\beta \leq 2(1 + |x|)^{-\beta'}.$$

We also note that  $\widehat{\Phi}_r(x) < C_{N_1}(r|x|)^{-N_1}$  for every  $N_1$  hence (6.35) is bounded from above by

$$(6.36) \quad 10 C_{d,N} c_\eta \int \left| \widehat{\rho}(x) \widehat{\Phi}_r(x) \right|^2 (1 + |x|)^{\alpha-d} d\eta(x) + C'(N, \beta', \alpha) + \\ + 10 C_{d,N} c_\eta C_{N_1} \int_{|x| > r^{-\beta/\beta'}} r^\beta (r|x|)^{-N_1} |x|^{\beta'+\alpha-d} d\eta.$$

As long as  $N_1$  is large enough (depending on  $\beta, \beta', d, \alpha$ ), the integral on the second line of (6.36) is bounded by a constant (depending on the same set of parameters).  $\square$

As in Section 6.B, we interpret the identity

$$\widehat{\mu}_n(b) = \sum_g \nu^{*m}(g) \widehat{\mu}_{n-m}(g^{\text{tr}}b)$$

to mean that for “many”  $g$  in the support of  $\nu^{*m}$ , the set of large Fourier coefficients  $A_{t',n-m}$  of  $\mu_{n-m}$  contains “a substantial proportion of”  $g^{\text{tr}}A$ . This later sets we consider as a perturbation of a rescaled and rotated orthogonal projection of  $A$  in the direction  $g$  expands the most (in the notations of Section 4.A, the direction perpendicular to  $H(g)$ ).

**Lemma 6.12.** *There are  $\epsilon_0, C, c_8 > 0$  (depending on  $\nu$ ) and an absolute constant  $q > 0$  so that if for some  $1/2 > t > 0$ ,  $1 \leq M < N$  with*

$$(6.37) \quad \log(N/M) > c_8 \log(1/t) \quad \text{and} \quad n \geq c_8 \log(N/M)$$

*it holds that*

$$\mathcal{N}(A_{t,n} \cap \mathbf{B}_{0,N}; M) > \left( \frac{N}{M} \right)^{d-\epsilon_0},$$

*then there are  $m, M', N'$  with  $M' \geq M$*

$$m \leq c_8 \log(N/M) \quad N' \leq N(N/M)^{c_8} \quad (N'/M') \geq (N/M)^{1/c_8}$$

*and  $\xi \in \mathbb{P}^{d-1}$  so that if  $R$  denotes the “rectangle”  $\mathbf{B}_{0,N'} \cap \text{Nbd}_{M'}(\xi)$  and  $t' = Ct^q$ .*

$$(6.38) \quad \mathcal{N}(A_{t',n-m} \cap R; M') > \frac{t'N'}{M'}.$$

*Proof.* Let  $\tau$  be as in Lemma 4.5, and set  $\epsilon_0 = \tau/3$ . Assume that for  $t, n, M, N$  as in the statement of Lemma 6.12 we have that

$$\mathcal{N}(A_{t,n} \cap \mathbf{B}_{0,N}; M) > \left( \frac{N}{M} \right)^{d-\epsilon_0}.$$

By Lemma 6.7 applied with  $\epsilon = \tau/30$  there is a  $N_1 \in (M, N)$  with  $\log(N_1/M) > \frac{1}{2} \log(N/M)$  so that  $A_{t_1,n} \cap \mathbf{B}_{0,N_1}$  contains a subset  $E$  which is  $(Ct^{-2}, d - 2\tau/3)$ -regular at scale  $M$ , where  $t_1 = t^2/4$  and  $C$  depends only on  $\tau$ . As before, we may

assume

$$(6.39) \quad \frac{1}{|E|} \left| \sum_{b \in E} \widehat{\mu}_n(b) \right| \geq \frac{t_1}{2}$$

since we may always choose a subset  $E_1 \subset E$  of cardinality  $\geq |E|/4$  on which the above inequality holds wake a to  $(Ct^{-2}, d - 2\tau/3)$ -regular (possibly for a slightly different  $C$ ).

Let  $m_1 = \kappa \log(N/M)$  (for a large constant  $\kappa$  to be determined later depending on  $\nu$ ), and set  $n_1 = n - m_1$ . For any  $g \in \text{supp}(\nu^{*m_1})$  set

$$E(g) = E \cap (g^{\text{tr}})^{-1} A_{\frac{t_1}{8}, n_1}$$

$$\mathcal{G}_{stat} = \left\{ g \in \text{supp} \nu^{*m_1} : |E(g)| > \frac{t_1}{8} |E| \right\}.$$

By (6.39), as  $\mu_n = \nu^{*m_1} * \mu_{n_1}$ ,

$$\frac{1}{|E|} \sum_g \nu^{*m_1}(g) \left| \sum_{b \in E} \widehat{\mu}_{n_1}(g^{\text{tr}} b) \right| \geq \frac{t_1}{2},$$

and it follows that for a set of  $g$  of  $\nu^{*m_1}$ -measure at least  $t_1/4$  one has

$$\sum_{b \in E} \widehat{\mu}_{n_1}(g^{\text{tr}} b) \geq \frac{t_1}{4} |E|.$$

By Chebyshev inequality any such  $g$  satisfies  $|E(g)| \geq t_1 |E|/8$ ; hence we conclude that

$$\nu^{*m_1}(\mathcal{G}_{stat}) \geq \frac{t_1}{4}.$$

Let  $\omega = (\lambda_1 - \lambda_2)/20$ , and set

$$\mathcal{G}_{len} = \left\{ g \in \Gamma : \left| \lambda_i - \frac{1}{m_1} \log \sigma_i(g) \right| < \omega \text{ for } i = 1, 2 \right\}.$$

By Theorem 4.3 and (6.37), if  $c_8$  is sufficiently large (depending on  $\nu$ ),

$$\nu^{*m_1}(\mathcal{G}_{len}) > 1 - t_1/8$$

hence  $\nu^{*m_1}(\mathcal{G}_{len} \cap \mathcal{G}_{stat}) \geq t_1/8$ . Let  $\mathcal{G} = \mathcal{G}_{stat} \cap \mathcal{G}_{len}$  and  $\eta$  be the probability measure on  $\mathbb{P}^{d-1}$  defined by

$$\eta(\Omega) = \frac{\nu^{*m_1} \{g \in \mathcal{G} : \theta(g) \in \Omega\}}{\nu^{*m_1}(\mathcal{G})}.$$

By Lemma 4.5, for any  $\xi \in \mathbb{P}^{d-1}$  and  $e^{-c_1 m_1} < r < r_0$  (with  $r_0, C, c_1, \tau$  as in that lemma)

$$\eta(\mathbf{B}_{\xi, r}) \leq 4t_1^{-1} r^\tau.$$

Applying Proposition 6.11 with  $\beta = \tau$ ,  $\beta' = \frac{5\tau}{6}$ ,  $\alpha = d - \frac{5\tau}{6}$  and  $\rho = \frac{1}{|E|} \sum_{b \in E} \delta_{b/N_1}$  we get

$$(6.40) \quad \int_{\xi} \left\| \frac{d(\rho_{\xi} * \Psi_r)}{dx} \right\|_2^2 d\eta(\xi) \leq C' t_1^{-1} \left[ C_d \int_{\mathbb{R}^d} |\widehat{\rho}(x)|^2 |\widehat{\Phi}_r(x)|^2 (1 + |x|)^{\alpha-d} dx + C(\alpha, \beta, \beta', d) \right].$$

Recall that  $\rho$  is  $(ct^{-2}, d - \frac{2}{3}\tau)$ -regular at scale  $M/N_1$ ; moreover if  $\kappa \geq c_1^{-1}$ , we have that  $r := M/N_1 \geq e^{-c_1 m_1}$ . It follows that

$$\begin{aligned} \int_{\mathbb{R}^d} |\widehat{\rho}(x)|^2 |\widehat{\Phi}_r(x)|^2 (1 + |x|)^{\alpha-d} dx &\asymp \mathcal{E}_{\alpha}(\rho \circ \Psi_r) && \text{(by (5.2))} \\ &\leq c'' t^{-2} = 8c'' t_1^{-1} && \text{(since } \alpha > d - 2\tau/3) \end{aligned}$$

with  $c', c''$  depending on  $\tau, \nu$ . Substituting into (6.40) we get

$$\int_{\xi} \left\| \frac{d(\rho_{\xi} * \Psi_r)}{dx} \right\|_2^2 d\eta(\xi) \leq c_* t_1^{-2}.$$

We conclude that there is a  $g_0 \in \mathcal{G}$  for which

$$(6.41) \quad \left\| \frac{d(\rho_{\xi_0} * \Psi_r)}{dx} \right\|_2^2 \leq c_* t_1^{-2} \quad \text{with } \xi_0 = \theta(g_0).$$

Set

$$\begin{aligned} M' &= \max(N_1 e^{\sigma_2(g_0)}, M e^{\sigma_1(g_0)}) \\ N' &= N_1 e^{\sigma_1(g_0)}. \end{aligned}$$

Since  $g_0 \in \mathcal{G}_{len}$  we have that

$$\log(N'/M') \geq \min(\log(N_1/M), (\lambda_1 - \lambda_2 - 2\omega)m_1) \gg \log(N/M)$$

(the implicit constant depending on  $\nu$ ). Also clearly  $M' \geq M$ . Since  $g_0 \in \mathcal{G}_{stat}$ , we have that  $|E(g_0)| > t_1 |E|/8$  hence  $\rho(\frac{1}{N_1} E(g_0)) \geq t_1/8$ . Let  $\pi_{\xi_0}$  denote the orthogonal projection to the direction  $\xi_0$  (considered as a map  $\mathbb{R}^d \rightarrow \mathbb{R}$ ). By Lemma 6.10 and (6.41) it follows that

$$(6.42) \quad \mathcal{N} \left( \pi_{\xi_0} \left( \frac{1}{N_1} E(g_0) \right); r' \right) \geq c_{**} (r')^{-1} t_1^4$$

where  $r' = M'/N' \geq r$ , and  $c_{**} = 2^{-8} c_*^{-1}$ . By definition of  $E(g_0)$ , we have that  $g_0^{\text{tr}}(E(g_0)) \subset A_{t_1/8, n_1}$ ; moreover for  $b \in \mathbb{B}_{0, N_1}$

$$\left\| g_0^{\text{tr}} b - e^{\sigma_1(g_0)} \pi_{\xi_0}(b) \theta(g_0) \right\| \leq N_1 e^{\sigma_2(g_0)} \leq M'.$$

In particular, setting  $\xi = \theta(g)$  and  $R$  the rectangle  $\mathbb{B}_{0, N'} \cap \text{Nbd}_{M'}(\xi)$ ,

$$(6.43) \quad g_0^{\text{tr}}(E(g_0)) \subset R \cap A_{t_1/8, n_1}$$

$$(6.44) \quad \mathcal{N}(g_0^{\text{tr}}(E(g_0)); M') \geq \mathcal{N}(\pi_{\xi_0}(E(g_0)); r')/2.$$

By (6.42), (6.43), and (6.44), keeping in mind that  $r' = M'/N'$ , the desired inequality (6.38) follows.  $\square$

Similarly to the proof of Proposition 6.3, Proposition 6.5 can easily be deduced from Lemma 6.12 using Lemma 6.9. Note that in the notations of Lemma 6.9,  $\left| \log \frac{N_1}{M_1} - \log \frac{N_2}{M_2} \right| \ll \log t_1$  with the implicit constant depending on  $\nu$ . We omit the details.

## 7. GRANULATED MEASURES

The goal of this section is to prove Proposition 3.1 and hence our main result Theorem A which follows easily from it. Assume that  $\mu_n = \nu^{*n} * \mu_0$  satisfies

$$|\widehat{\mu}_{n_0}(a_0)| \geq t_0 > 0$$

where  $n_0$  is assumed to be larger than a constant multiple of  $\log(2\|a_0\|/t_0)$ . The goal is to deduce that for any  $\lambda < \lambda_1$  there is  $C$  so that

$$\mu_{n_0-m}(W_{Q,e^{-\lambda \cdot m}}) > \left(\frac{t_0}{2}\right)^C, \quad \text{where} \quad Q < \left(\frac{2\|a_0\|}{t_0}\right)^C.$$

We recall the notations

$$R_Q = \left\{ \left( \frac{p_1}{q}, \dots, \frac{p_d}{q} \right) \in \mathbb{T}^d : q \leq Q \right\}, \quad W_{Q,r} = \bigcup_{x \in R_Q} B_{x,r}.$$

Unless otherwise specified, *all other constants defined in this section depend only on  $\nu$*  (and hence indirectly also on  $\Gamma$ ).

We outline the ingredients of the argument in the following Propositions 7.1–7.4, and formally deduce Proposition 3.1. The proofs of propositions 7.1–7.4 are given in §§7.A–7.D below.

In the first phase of the proof (Section 6, Theorem 6.1) it was shown that the set of significant Fourier coefficients  $\{a \in \mathbb{Z}^d : |\widehat{\mu}_{n_0-m}(a)| > t\}$  in large balls  $\{a \in \mathbb{Z}^d : \|a\| < N\}$  has positive density when viewed at resolution  $M = N^{1-\kappa}$ . We shall use this information on Fourier coefficients to show that a certain portion of the measure  $\mu_{n_0-m}$  on the torus  $\mathbb{T}^d$  is  $(1-\kappa)$ -granulated at scale  $\rho = 1/N$  in the following sense.

Let  $\mu$  be a probability measure  $\mu$  on  $\mathbb{T}^d$ . Say that a  $t$ -portion of  $\mu$  is  $\alpha$ -granulated at scale  $\rho$  (here  $\alpha < 1$  and  $\rho > 0$  is smaller than a power of  $t/2$ ) if there exists a  $\rho^\alpha$ -separated set  $X \subset \mathbb{T}^d$  so that

$$\mu(\text{Nbd}_\rho(X)) = \mu\left(\bigcup_{x \in X} B_{x,\rho}\right) > t.$$

The information on significant Fourier coefficients of  $\mu_n$  obtained in the first phase of the proof (Section 6, Theorem 6.1) enables one to show that a significant portion of the measures  $\mu_{n_0-m}$  is  $(1-\kappa)$ -granulated.

**Proposition 7.1** (Initial granulation estimate).

*There exist constants  $1 < L_1 < L_2$ ,  $\kappa > 0$ , and  $c_1, c_2$ , so that for  $m \geq c_1 \cdot \log \frac{2\|a_0\|}{t_0}$ , there exists  $\rho \in (L_2^{-m}, L_1^{-m})$  and a finite set  $X \subset \mathbb{T}^d$  so that*

- (1)  $X$  is  $r = \rho^{1-\kappa}$ -separated;  
 (2)  $\mu_{n_0-m}(\bigcup_{x \in X} \mathbf{B}_{x,\rho}) > (\frac{t_0}{2})^{c_2}$ .

Let us say that a probability measure  $\mu$  is  $\beta$ -concentrated around  $x \in \mathbb{T}^d$  at scale  $\rho$  if  $\mu(\mathbf{B}_{x,\rho}) > \rho^\beta$ . So Lebesgue measure is  $d$ -concentrated, while atomic measures are 0-concentrated, at all scales. Observe that if  $\alpha < d$  and  $\alpha \cdot d < \beta < d$  then a probability measure  $\mu$  which is  $\alpha$ -granulated at sufficiently small scale  $\rho$  has points which are  $\beta$ -concentrated: since  $\rho^\alpha$ -separated subset on the  $d$ -torus has  $O(\rho^{-d\alpha})$  points, an average  $\rho$ -ball with center  $x \in X$  has  $\mu$ -mass

$$\mu(\mathbf{B}_{x,\rho}) > \text{const} \cdot t \cdot \rho^{\alpha \cdot d} > \rho^\beta.$$

Thus  $\mu_{n_0-m}$  has points which are  $\beta$ -concentrated where  $\beta = d - \kappa > (1 - \kappa) \cdot d$ , assuming the scale  $\rho$  is small compared to  $t$ . The next step of the argument allows to bootstrap this concentration phenomenon from  $\beta_0 = d - \kappa$  down to  $\beta_N = \delta$ , where  $\delta > 0$  is some fixed concentration goal determined in Proposition 7.3 below. The bootstrapping procedure is performed some finite number  $N = N(\kappa, \delta)$  of times.

**Proposition 7.2** (Bootstrapping concentration).

Given  $\epsilon > 0$  there is  $\gamma > 0$  and  $\ell_0$  so that for  $n > \ell > \ell_0$  the following holds: given scales  $\rho < e^{-d\lambda_1 \cdot \ell} \cdot r$  there are scales

$$r' = e^{-(\lambda_1 + \epsilon) \cdot \ell} \cdot r, \quad \rho' = e^{-(\lambda_1 - \epsilon) \cdot \ell} \cdot \rho$$

so that given an  $r$ -separated set  $X \subset \mathbb{T}^d$  one can construct an  $r'$ -separated set  $X' \subset \mathbb{T}^d$  with

$$|X'| \leq |X| \quad \text{and} \quad \mu_{n-\ell}(\bigcup_{y \in X'} \mathbf{B}_{y,\rho'}) > \left( \mu_n(\bigcup_{x \in X} \mathbf{B}_{x,\rho}) \right)^d - e^{-\gamma \cdot \ell}.$$

The initial granulation  $\alpha = 1 - \kappa$  gives  $\frac{r_0}{\rho_0} = \rho_0^{-\kappa}$  so the above proposition can be applied with  $\ell$  as big as  $\frac{1}{d\lambda_1} \log(\frac{r_0}{\rho_0}) = \frac{\kappa}{d\lambda_1} \log(\frac{1}{\rho_0})$ . With half that big  $\ell$ , we still get a shrinking factor of  $e^{-(\lambda_1 - \epsilon) \cdot \ell} < \rho_0^{\kappa/3d}$  in the scale of concentrated balls produced in the proposition. The fact that the ratio  $\frac{r'}{\rho'}$  in the output is close to the initial one  $\frac{r}{\rho}$ , allows to apply the proposition with a fixed  $\ell$  for a number  $N$  of iterations, and obtain very high concentrations. The loss of mass is not very drastic if the initial portion  $\tau_0 > (t_0/2)^{c_2}$  of  $(1 - \kappa)$ -granulated measure  $\mu_{n_0-m}$  is large compared to the scale  $\rho$  and  $e^{-\gamma \ell}$ .

The following proposition shows that certain level of concentration can occur only near rational points. This determines the desired concentration level  $\delta > 0$  mentioned above.

**Proposition 7.3** (Rational approximation).

There is  $\delta > 0$  and  $c_4 < \infty$  so that for any small  $\rho > 0$

$$\mu_n(\mathbf{B}_{z,\rho}) > \rho^\delta \quad \implies \quad \mathbf{B}_{z,\rho} \subset W_{Q,r},$$

for  $r = \rho^{9/10}$  and  $Q = \rho^{-1/10}$ , provided  $n > c_4 \cdot \log(1/\rho)$ .

Hence assuming that a significant  $\mu_n$ -mass is granulated with exponent  $\delta$ ,  $\mu_n$  gives this significant mass to  $W_{Q,r}$  with  $r = Q^{-9}$ . Of course factor 9 is arbitrary here, for the following we could work with any factor bigger than say 3.

The final step of the proof, uses the  $\Gamma$ -invariance of the set  $R_Q$  to show that most of  $\mu_n$  mass of  $W_{Q,r} = \bigcup_{x \in R_Q} B_{x,r}$  must be concentrated near the centers  $R_Q$  of these balls.

**Proposition 7.4** (Tight Bootstrapping).

Given  $\epsilon > 0$  there is  $m_*$  and  $\omega > 0$  so that if  $r > 0$ ,  $Q < \infty$  and  $m > m_*$  satisfy

$$e^{d\lambda_1 \cdot m} \cdot r < \frac{1}{Q^2}$$

then

$$\mu_{n-m}(W_{Q,e^{-(\lambda_1-\epsilon) \cdot m} \cdot r}) > \mu_n(W_{Q,r}) - e^{-\omega \cdot m}$$

assuming  $n > m$ .

This is done by considering the intersections of a large number  $N > e^{\delta \cdot m}$  of translates  $g_i^{-1}(W_{Q,r})$ , where  $g_1, \dots, g_N$  are chosen using the distribution  $\nu^{*m}$  of the  $m$ -step random walk.

Let us now deduce Proposition 3.1 of the introduction from these propositions, which are proved in the following subsections §§7.A–7.D.

*Proof of Proposition 3.1.*

We assume that  $|\widehat{\mu}_{n_0}(a_0)| \geq t_0 > 0$  for some  $a_0 \in \mathbb{Z}^d \setminus \{0\}$ . We shall work with  $n_0 > m > C \cdot \log \frac{2\|a_0\|}{t_0}$  where the value of  $C$  will be determined implicitly in the proof.

Our first goal is to show that for some constants  $C_1, D$ ,  $1 < L_3 < L_4$  and any  $m_0 > C_1 \cdot \log \frac{2\|a_0\|}{t_0}$  there exist  $\rho$  with  $L_4^{-m_0} < \rho < L_3^{-m_0}$  and a finite set  $Y \subset \mathbb{T}^d$  so that

$$(7.1) \quad \mu_{n_0-m_0}(B_{y,\rho}) > \rho^\delta \quad (\forall y \in Y), \quad \mu_{n_0-m_0}\left(\bigcup_{y \in Y} B_{y,\rho}\right) > \left(\frac{t_0}{2}\right)^D$$

where  $\delta > 0$  is the constant from Proposition 7.3.

Proposition 7.1 provides  $1 < L_1 < L_2$  and  $\kappa > 0$ , so that for large  $m_{00}$  there exists  $\rho_0 \in (L_2^{-m_{00}}, L_1^{-m_{00}})$  and finite set  $X_0 \subset \mathbb{T}^d$  which is  $r_0$ -separated so that

$$r_0 = \rho_0^{1-\kappa}, \quad \mu_{n_0-m_{00}}\left(\bigcup_{x \in X_0} B_{x,\rho_0}\right) > \left(\frac{t_0}{2}\right)^{c_2}.$$

We shall amplify this initial concentration by a number ( $N$  below) of iterations of the bootstrapping procedure in Proposition 7.2. The relevant parameters are

chosen as follows:

$$(7.2) \quad \ell \in \mathbb{N} \quad \text{so that} \quad e^{-2d\lambda_1 \cdot \ell} > \frac{\rho_0}{r_0} = \rho_0^\kappa > e^{-3d\lambda_1 \cdot \ell},$$

$$(7.3) \quad N \in \mathbb{N} \quad \text{so that} \quad \delta N \cdot \kappa > 6d^2,$$

$$(7.4) \quad \epsilon > 0 \quad \text{so that} \quad 2N \cdot \epsilon < d\lambda_1.$$

Here  $\delta > 0$  is provided by Proposition 7.3 and  $\kappa$  by Proposition 7.1. Note that  $\ell \asymp \log \frac{1}{\rho_0} \asymp m_{00}$ , i.e., the ratios between these quantities are bounded from below and from above by finite positive constants (depending on  $\nu$ ).

For  $j = 1, \dots, N-1$  set  $\rho_{j+1} = e^{-j(\lambda_1 - \epsilon) \cdot \ell} \cdot \rho_0$  and  $r_{j+1} = e^{-j(\lambda_1 + \epsilon) \cdot \ell} \cdot r_0$ . Then

$$(7.5) \quad \frac{\rho_0}{r_0} < \dots < \frac{\rho_N}{r_N} = e^{2N\epsilon\ell} \cdot \frac{\rho_0}{r_0} < e^{2N\epsilon\ell - 2d\lambda_1\ell} < e^{-d\lambda_1\ell},$$

where the last inequality is justified by (7.2) and (7.4).

We have arranged  $\rho_j < e^{-d\lambda_1\ell} \cdot r_j$  for  $j = 0, \dots, N$ , and, assuming that  $\ell > \ell_0$ , may apply Proposition 7.2 inductively starting from the set  $X_0$  provided by Proposition 7.1. This yields a finite sequence of sets  $X_1, \dots, X_N$ , where each  $X_j$  is an  $r_j$ -separated set on the torus, the sets do not increase in cardinality:

$$(7.6) \quad |X_N| \leq \dots \leq |X_1| \leq |X_0| < \text{const}_d \cdot r_0^{-d} < \rho_0^{-d},$$

while the masses

$$\tau_j = \mu_{n_0 - j\ell} \left( \bigcup_{x \in X_j} \mathbf{B}_{x, \rho_j} \right) \quad \text{satisfy} \quad \tau_{j+1} > \tau_j^d - e^{-\gamma \cdot \ell}.$$

Recall that  $\gamma > 0$ , depending on  $\epsilon > 0$ ,  $N$ ,  $\kappa > 0$  and  $\delta > 0$ , are constants (depending on  $\nu$ ), but independent of  $\ell$ ,  $m_{00}$ , etc. So choosing  $C_1$  large enough, we may ensure that  $m_{00}$ , and thus  $\ell$ , are large compared to  $\log(2/t_0)$  so that

$$e^{-\gamma \cdot \ell} < \left( \frac{t_0}{2} \right)^{c_2 \cdot (d+1)^N}.$$

This implies, by induction on  $i$ , that  $\tau_i > 2e^{-\gamma \cdot \ell}$  and  $\tau_{i+1} > \frac{1}{2} \tau_i^d > \tau_i^{d+1}$ . In particular the last set  $X_N$  satisfies:

$$\mu_{n_0 - N\ell} \left( \bigcup_{x \in X_N} \mathbf{B}_{x, \rho_N} \right) = \tau_N > \left( \frac{t_0}{2} \right)^{c_2 \cdot (d+1)^N}.$$

We now use the fact that  $|X_N|$  has few elements, estimated by (7.6), to extract the subset  $Y$  of very concentrated  $\rho_N$ -balls:

$$(7.7) \quad Y = \left\{ x \in X_N : \mu_{n_0 - N\ell}(\mathbf{B}_{x, \rho_N}) > \frac{\tau_N}{2 \cdot |X_N|} \right\}.$$

Then

$$\mu_{n_0 - N\ell} \left( \bigcup_{y \in Y} \mathbf{B}_{y, \rho_N} \right) > \frac{\tau_N}{2} > \left( \frac{t_0}{2} \right)^D,$$

where  $D$  is set to be  $D = c_2 \cdot (d+1)^N + 1$ . Finally we claim that

$$(7.8) \quad \frac{\tau_N}{2|X_N|} > (\rho_N)^\delta.$$

Indeed, assuming  $m_{00}$  is large compared to  $\log(2/t_0)$ , we have

$$\frac{\tau_N}{2|X_N|} > \left(\frac{t_0}{2}\right)^D \cdot \rho_0^{d(1-\kappa)} > \rho_0^d.$$

Using (7.3) and (7.2) it follows

$$(\rho_N)^\delta = e^{-\delta N(\lambda_1 - \epsilon) \cdot \ell} \cdot \rho_0^\delta < e^{-\delta N \frac{\lambda_1}{2} \ell} < \left(e^{-3d\lambda_1 \ell}\right)^{d/\kappa} < \rho_0^d.$$

With  $Y$  as in (7.7),  $\rho = \rho_N$ ,  $m_0 = N\ell$ , the claim (7.1) is proven.

Applying Proposition 7.3 to the conclusion (7.1) we deduce that for some  $C_2, C_3 > 1$ , for  $m_0 > C_2 \cdot \log \frac{2\|a_0\|}{t_0}$ , and  $n_0 > C_3 \cdot m_0$ , one has

$$(7.9) \quad \mu_{n_0 - m_0}(\mathbb{W}_{Q,r}) > \left(\frac{t_0}{2}\right)^D, \quad \text{where} \quad r = Q^{-9}, \quad Q \in (L_3^{\frac{m_0}{10}}, L_4^{\frac{m_0}{10}}).$$

The proof of Proposition 3.1 concludes with the second bootstrap Proposition 7.4 applied a number of times. Given  $\lambda < \lambda_1$  we choose

$$\epsilon = \min\left(\frac{\lambda_1}{3}, \frac{\lambda_1 - \lambda}{2}\right)$$

and let  $\omega = \omega(\epsilon) > 0$  be the corresponding constant from Proposition 7.4.

With  $\epsilon < \lambda_1 - \lambda$  there are  $0 < \alpha < \beta$  and  $k \in \mathbb{N}$ , so that any large  $m$  can be written as

$$m = m_0 + m_1 + m_2 + \cdots + m_k,$$

where

$$(7.10) \quad \lambda m < (\lambda_1 - \epsilon) \cdot (m - m_0),$$

$$(7.11) \quad \alpha \cdot m < m_0 < \beta \cdot m,$$

$$(7.12) \quad \left(\frac{7}{10d\lambda_1} \log L_3\right) \cdot m_0 < m_1 < \left(\frac{7}{10d\lambda_1} \log L_4\right) \cdot m_0,$$

$$(7.13) \quad \left(1 + \frac{1}{3d}\right) \cdot m_i < m_{i+1} < \left(1 + \frac{1}{2d}\right) \cdot m_i.$$

We set  $C$  to be large enough so that writing  $m > C \cdot \log \frac{2\|a_0\|}{t_0}$  as  $m = m_0 + \cdots + m_k$  as above we get  $m_0 > C_2 \cdot \log \frac{2\|a_0\|}{t_0}$  and  $m_1 > m_*$ . Then for  $r$  and  $Q$  as in (7.9) condition (7.12) implies

$$e^{d\lambda_1 m_1} < L_3^{\frac{7m_0}{10}} < Q^7 = \frac{1}{r \cdot Q^2}.$$

Denoting  $r_i = e^{-(\lambda_1 - \epsilon) \cdot (m_1 + \cdots + m_i)} \cdot r$  we also obtain

$$e^{d\lambda_1 \cdot m_{i+1}} \cdot r_i < \frac{1}{Q^2}.$$



Indeed, this is proven by induction using (7.13):

$$e^{d\lambda_1 \cdot m_{i+1}} < e^{d\lambda_1 \cdot m_i} \cdot e^{\frac{\lambda_1}{2} \cdot m_i} < \frac{e^{\frac{\lambda_1}{2} \cdot m_i}}{r_i \cdot Q^2} < \frac{1}{r_{i+1} \cdot Q^2}.$$

Therefore, Proposition 7.4 can be applied to deduce, using (7.10), that

$$\begin{aligned} \mu_{n_0-m}(\mathbb{W}_{Q,e^{-\lambda \cdot m}}) &> \mu_{n_0-m}(\mathbb{W}_{Q,e^{-(\lambda_1-\epsilon) \cdot (m_1+\dots+m_k)} \cdot r}) \\ &> \mu_{n_0-m_0}(\mathbb{W}_{Q,r}) - e^{-\omega \cdot m_1} - \dots - e^{-\omega \cdot m_k}. \end{aligned}$$

For some  $c > 0$ , independent of  $m$ , etc, we have  $\sum e^{-\omega \cdot m_i} < e^{-c \cdot m}$ . If  $C > 2D/c$  then it follows, using (7.9), that

$$\mu_{n_0-m}(\mathbb{W}_{Q,e^{-\lambda \cdot m}}) > \mu_{n_0-m_0}(\mathbb{W}_{Q,r}) - e^{-c \cdot m} > \left(\frac{t_0}{2}\right)^D - e^{-c \cdot m} > \left(\frac{t_0}{2}\right)^{D+1}.$$

This completes the proof of Proposition 3.1.  $\square$

#### 7.A. Initial granulation: proof of Proposition 7.1.

Proposition 7.1 follows from Theorem 6.1 and the following general statement with  $M = N^{1-\kappa}$ ,  $\rho = \frac{1}{M}$ ,  $s = t = t_0^p$ .

**Proposition 7.5.** *There exists  $c > 0$  so that if a probability measure  $\mu$  on  $\mathbb{T}^d$  satisfies*

$$\mathcal{N}\left(\left\{a \in \mathbb{Z}^d \cap \mathbb{B}_{0,N} : |\hat{\mu}(a)| > t\right\}; M\right) > s \cdot \left(\frac{N}{M}\right)^d$$

with  $M < \text{const}_d \cdot N$ , then there exists an  $\frac{1}{M}$ -separated set  $X \subset \mathbb{T}^d$  with

$$\mu\left(\bigcup_{x \in X} \mathbb{B}_{x, \frac{1}{N}}\right) > c \cdot (ts)^3.$$

*Proof.* We shall need an auxiliary smooth function  $F$  on the torus such that

$$0 \leq F \leq C_1 \cdot N^d, \quad \text{supp}(F) \subset \mathbb{B}_{0, \frac{1}{N}}, \quad \int_{\mathbb{T}^d} F dx = 1.$$

and the Fourier coefficients

$$\hat{F}(a) \geq 0 \quad \hat{F}(a) \geq \frac{1}{2} \quad \text{for } a \in \mathbb{Z}^d \cap \mathbb{B}_{0,N}.$$

Here  $C_1 < \infty$  is a constant depending on  $d$  only. To construct such a function, consider the step function  $F_1(x) = m(\mathbb{B}_{0,r})^{-1} \cdot 1_{\mathbb{B}_{0,r}}(x)$  where  $r = \epsilon/N$  for some fixed small  $\epsilon > 0$ . Then  $\hat{F}_1(a)$  is close to 1 for  $a \in \mathbb{Z}^d \cap \mathbb{B}_{0,N}$ . If  $F_2$  is a smooth symmetric approximation of  $F_1$ , then the convolution  $F = F_2 * \check{F}_2$  has the desired properties.

Let  $\tilde{A} = \{a \in \mathbb{Z}^d \cap \mathbb{B}_{0,N} : |\hat{\mu}(a)| > t\}$ . Upon passing to a subset  $A \subset \tilde{A}$  of size

$$|A| \geq \frac{|\tilde{A}|}{4} > \frac{s}{4} \left(\frac{N}{M}\right)^d$$

we may assume that  $\operatorname{Re}(e^{i\theta} \cdot \hat{\mu}(a)) > \frac{t}{2}$  for some fixed  $\theta \in [0, 2\pi)$ . Let

$$\phi(x) = \sum_{a \in A} e_a(x).$$

As usual,  $e_a(x) = e^{-2\pi i \langle x, a \rangle}$  are the standard characters. Note that

$$|\phi(x)|^2 = \left( \sum_{a \in A} e_a(x) \right) \cdot \overline{\left( \sum_{b \in A} e_b(x) \right)} = \sum_{a, b \in A} e_{a-b}(x).$$

The probability measure  $\lambda = \mu * F$  has a smooth density  $g : \mathbb{T}^d \rightarrow [0, \infty)$  with  $\hat{g}(b) = \hat{\mu}(b) \cdot \hat{F}(b)$ . On  $A$  we have  $\hat{F} \geq 1/2$  and  $\operatorname{Re}(e^{i\theta} \hat{\mu}) > t/2$ . Therefore

$$(7.14) \quad \left| \int_{\mathbb{T}^d} \phi \, d\lambda \right| \geq \sum_{a \in A} \operatorname{Re}(e^{i\theta} \cdot \hat{g}(a)) > \frac{t}{4} \cdot |A| > \frac{ts}{16} \cdot \left( \frac{N}{M} \right)^d.$$

We shall see that the RHS is close to an a priori upper estimate for the LHS. Partition  $\mathbb{T}^d$  into  $M^d$  "cubes"  $Q_i$  with side length  $\frac{1}{M}$  and centers  $c_i \in \mathbb{T}^d$ . By Cauchy-Schwartz inequality

$$(7.15) \quad \left| \int_{\mathbb{T}^d} \phi \, d\lambda \right| \leq \sum_i \left| \int_{\mathbb{T}^d} 1_{Q_i} \cdot \phi \, d\lambda \right| \leq \sum_i \lambda(Q_i)^{\frac{1}{2}} \cdot \left( \int_{Q_i} |\phi|^2 \, d\lambda \right)^{\frac{1}{2}}.$$

Let  $r = \frac{\sqrt{d}}{M}$  which is assumed to dominate  $\frac{1}{N}$ . Then  $Q_i \subset \mathbf{B}_{c_i, r/2}$  and  $y + Q_i \subset \mathbf{B}_{c_i, r}$  for any  $y \in \operatorname{supp}(F) \subset \mathbf{B}_{0, \frac{1}{N}}$ . Thus

$$\lambda(Q_i) = \int_{\mathbb{T}^d} F(y) \cdot \mu(y + Q_i) \, dy \leq \mu(\mathbf{B}_{c_i, r}).$$

Since  $d\lambda(x) = g(x) \, dx$  we have

$$\int_{Q_i} |\phi|^2 \, d\lambda \leq G_i \cdot \int_{Q_i} |\phi|^2 \, dx, \quad \text{where} \quad G_i = \max_{x \in Q_i} g(x).$$

We shall estimate  $\int_{Q_i} |\phi|^2 \, dx$  using an auxiliary function  $f$  on  $\mathbb{T}^d$ ; we take  $f$  to be the product  $f(x) = \prod_{i=1}^d h_M(x_i)$  of one dimensional Fejér kernels

$$h_n(u) = \frac{1}{n} \sum_{k=1}^n \sum_{j=-k}^k e^{2\pi j u} = \frac{1}{n} \left( \frac{\sin \frac{nu}{2}}{\sin \frac{u}{2}} \right)^2.$$

Note that  $f$  is a non-negative function, with  $f(x) > 10^{-d} \cdot M^d$  on the  $\frac{1}{M}$ -cube  $Q_0 = [-\frac{1}{2M}, \frac{1}{2M}]^d + \mathbb{Z}^d$  around  $0 \in \mathbb{T}^d$ . The Fourier coefficients  $\hat{f}$  take values in

$[0, 1]$  and vanish outside the  $[-M, M]^d \cap \mathbb{Z}^d$  cube. Thus

$$\begin{aligned} \int_{Q_i} |\phi(x)|^2 dx &= \int_{Q_0} |\phi(c_i + y)|^2 dy \leq \frac{10^d}{M^d} \int_{Q_0} |\phi(c_i + y)|^2 f(y) dy \\ &\leq \frac{10^d}{M^d} \int_{\mathbb{T}^d} |\phi(c_i + y)|^2 f(y) dy = \frac{10^d}{M^d} \int_{\mathbb{T}^d} \sum_{a,b \in A} e_{a-b}(c_i + y) \cdot f(y) dy \\ &= \frac{10^d}{M^d} \left( \sum_{a,b \in A} e_{a-b}(c_i) \hat{f}(a-b) \right) \leq \frac{10^d}{M^d} \cdot \max_{a,b \in A} |\hat{f}(a-b)| \\ &\leq \frac{C_2 \cdot |A|}{M^d} \leq \frac{C_2 \cdot N^d}{M^{2d}}. \end{aligned}$$

Here  $C_2$  is a constant which is  $10^d$  times the maximal cardinality of a 1-separated set in  $[-1, 1]^d$ .

The density  $g$  of  $\lambda = \mu * F$  has the following upper bound:

$$(7.16) \quad g(x) = \int F(x-y) d\mu(y) \leq C_1 \cdot N^d \cdot \mu(\mathbf{B}_{x, \frac{1}{N}}).$$

Since  $\text{Nbd}_{\frac{1}{N}}(Q_i) \subset \mathbf{B}_{c_i, r}$  it follows that

$$G_i = \max_{x \in Q_i} g(x) \leq C_1 N^d \mu(\mathbf{B}_{c_i, r}).$$

Let  $0 \leq H_i \leq 1$  denote the ratio, so  $G_i = H_i \cdot C_1 N^d \mu(\mathbf{B}_{c_i, r})$ . By (7.14 and 7.15)

$$\begin{aligned} \frac{ts}{16} \left( \frac{N}{M} \right)^d &\leq \sum_i \mu(\mathbf{B}_{c_i, r})^{\frac{1}{2}} \cdot G_i^{\frac{1}{2}} \cdot \frac{\sqrt{C_2} \cdot N^{\frac{d}{2}}}{M^d} \\ &\leq \sum_i \mu(\mathbf{B}_{c_i, r}) \cdot H_i^{\frac{1}{2}} \cdot \sqrt{C_1 \cdot C_2} \cdot \left( \frac{N}{M} \right)^d. \end{aligned}$$

Let  $C_3 = \sqrt{C_1 \cdot C_2}$ . We have

$$\sum_i \mu(\mathbf{B}_{c_i, r}) \cdot H_i^{\frac{1}{2}} > \frac{ts}{16C_3}.$$

Therefore

$$(7.17) \quad \sum_{i \in I} \mu(\mathbf{B}_{c_i, r}) > \frac{ts}{2^5 C_3} \quad \text{where} \quad I = \left\{ i : H_i^{\frac{1}{2}} > \frac{ts}{2^5 C_3} \right\}.$$

For each  $i \in I$  let  $x_i \in Q_i$  be such that  $g(x_i) > \left( \frac{ts}{2^5 C_3} \right)^2 \cdot C_1 \cdot N^d \cdot \mu(\mathbf{B}_{c_i, r})$ . Then

$$\mu(\mathbf{B}_{x_i, \frac{1}{N}}) > \frac{(ts)^2}{2^{10} C_2} \cdot \mu(\mathbf{B}_{c_i, r})$$

and using (7.17)

$$\sum_{i \in I} \mu(\mathbf{B}_{x_i, \frac{1}{N}}) > \frac{(ts)^3}{2^{15} \cdot C_2 \cdot C_3}.$$

The set  $\tilde{X} = \{x_i : i \in I\}$  visits each of the cubes  $Q_j$  at most ones. Thus it may be separated into  $2^d$  subsets each of which never visits neighboring  $Q_j$ -s, and is therefore  $\frac{1}{M}$ -separated. At least one of  $2^d$  such subsets  $X \subset \tilde{X}$  has

$$\mu\left(\bigcup_{x \in X} B_{x,r}\right) = \sum_{x \in X} \mu(B_{x,\frac{1}{N}}) > 2^{-d} \cdot \sum_{i \in I} \mu(B_{x_i,\frac{1}{N}}) > \frac{(ts)^3}{2^{d+15} \cdot C_2 \cdot C_3}.$$

This completes the proof of the Proposition.  $\square$

### 7.B. Bootstrapping the concentration: proof of Proposition 7.2.

We start with a few lemmas.

**Lemma 7.6.** *Given  $\epsilon > 0$  there is  $\gamma > 0$  and  $m_0 \in \mathbb{N}$  so that for  $n > m \geq m_0$  one can find a subset  $\mathcal{G} \subset \Gamma^d$  so that for  $(g_1, \dots, g_d) \in \mathcal{G}$ :*

- (i)  $\left| \frac{1}{m} \log \sigma_j(g_i) - \lambda_j \right| < \epsilon \quad (1 \leq i \leq d, 1 \leq j \leq d)$
- (ii)  $\text{vol}(\theta(g_1), \dots, \theta(g_d)) > e^{-\epsilon \cdot m}$
- (iii)  $\text{vol}(\theta(g_1^{\text{tr}}), \dots, \theta(g_d^{\text{tr}})) > e^{-\epsilon \cdot m}$

and such that for any Borel subset  $A \subset \mathbb{T}^d$  one has

$$\mu_n(A)^d - e^{-\gamma \cdot m} \leq \sum_{\vec{g} \in \mathcal{G}} \nu^{*m}(g_1) \cdots \nu^{*m}(g_d) \cdot \mu_{n-m}(g_1^{-1}A \cap \cdots \cap g_d^{-1}A).$$

*Proof.* By Theorem 4.3 for some  $\rho > 0$  and sufficiently large  $m$  the set  $\mathcal{G}_{\text{len}}$  of  $d$ -tuples  $\vec{g} \in \Gamma^d$  satisfying (i) has

$$(\nu^{*m})^d(\mathcal{G}_{\text{len}}) > (1 - e^{-\rho \cdot m})^d.$$

The set  $\mathcal{G}_{\text{trans}}$  of sufficiently "transversal"  $d$ -tuples  $\vec{g} \in \Gamma^d$ , namely ones satisfying conditions (ii) and (iii), has (Lemma 4.6) mass

$$(\nu^{*m})^d(\mathcal{G}_{\text{trans}}) > 1 - e^{-(\epsilon/p) \cdot m}.$$

Let  $\mathcal{G} = \mathcal{G}_{\text{len}} \cap \mathcal{G}_{\text{trans}}$  and let  $\gamma > 0$  be small enough so that

$$(\nu^{*m})^d(\mathcal{G}) > (1 - e^{-\rho \cdot m})^d - e^{-(\epsilon/p) \cdot m} > 1 - e^{-\gamma \cdot m}.$$

Given  $A \subset \mathbb{T}^d$  the function  $f(x) = \sum_{g \in \Gamma} \nu^{*m}(g) \cdot 1_A(gx)$  on  $\mathbb{T}^d$  satisfies

$$\int_{\mathbb{T}^d} f(x) d\mu_{n-m}(x) = \sum_{g \in \Gamma} \nu^{*m}(g) \cdot \mu_{n-m}(g^{-1}A) = \mu_n(A).$$

By the convexity of  $t \mapsto t^d$  we deduce that

$$\begin{aligned} \mu_n(A)^d &= \left( \int_{\mathbb{T}^d} f d\mu_{n-m} \right)^d \leq \int_{\mathbb{T}^d} f(x)^d d\mu_{n-m}(x) \\ &= \sum_{\vec{g} \in \Gamma^d} \nu^{*m}(g_1) \cdots \nu^{*m}(g_d) \cdot \mu_{n-m}(g_1^{-1}A \cap \cdots \cap g_d^{-1}A) \end{aligned}$$

and the Lemma follows by restricting the summation to  $\vec{g} \in \mathcal{G}$ .  $\square$

**Lemma 7.7.** *For any  $\bar{x}_1, \dots, \bar{x}_d, \bar{y}_1, \dots, \bar{y}_d \in \mathbb{P}^{d-1}$  one has*

$$|\text{vol}(\bar{x}_1, \dots, \bar{x}_d) - \text{vol}(\bar{y}_1, \dots, \bar{y}_d)| \leq \sqrt{2} \cdot \sum_{i=1}^d d_{\text{ang}}(\bar{x}_i, \bar{y}_i).$$

*Proof.* Assuming  $x_i, y_i$  are unit vectors, we have

$$\begin{aligned} |\text{vol}(x_1, \dots, x_d) - \text{vol}(y_1, \dots, y_d)| &\leq \sum_{i=1}^d |\text{vol}(x_1, \dots, x_i - y_i, \dots, y_d)| \\ &\leq \sum_{i=1}^d \|x_i - y_i\| \leq \sqrt{2} \cdot \sum_{i=1}^d d_{\text{ang}}(\bar{x}_i, \bar{y}_i). \end{aligned}$$

□

**Lemma 7.8.** *Given  $\epsilon > 0$  there is  $m_0(\epsilon)$  so that for  $m > m_0$  and any  $g_1, \dots, g_d \in \Gamma$  with*

$$\begin{aligned} \left| \frac{1}{m} \sigma_j(g) - \lambda_j \right| &< \epsilon, \quad (j = 1, 2) \\ \text{vol}(\theta(g_1^{\text{tr}}), \dots, \theta(g_d^{\text{tr}})) &> e^{-\epsilon \cdot m} \end{aligned}$$

*one has*

$$\forall v \in \mathbb{R}^d \setminus \{0\} : \quad \max_{1 \leq i \leq d} \frac{\|g_i v\|}{\|v\|} \geq e^{(\lambda_1 - 3\epsilon) \cdot m}.$$

*Proof.* First let us estimate

$$\delta = \max_{1 \leq i \leq d} d_{\text{ang}}(v, H(g_i)) = \max_{1 \leq i \leq d} d_{\text{ang}}(\theta(g_i^{\text{tr}}), v^\perp).$$

If  $y_i$  denote the projections of some unit vectors in  $\bar{x}_i = \theta(g_i^{\text{tr}})$  to  $v^\perp$ , then  $\text{vol}(\bar{y}_1, \dots, \bar{y}_d) = 0$ . Hence it follows from Lemma 7.7 that

$$\sqrt{2} \cdot \sum_{i=1}^d d_{\text{ang}}(\bar{x}_i, \bar{y}_i) \geq \text{vol}(\bar{x}_1, \dots, \bar{x}_d) > e^{-\epsilon \cdot m}.$$

Thus  $\delta > \frac{1}{\sqrt{2d}} \cdot e^{-\epsilon \cdot m}$ , which is larger than  $e^{-2\epsilon \cdot m}$  for sufficiently large  $m$ . We have

$$\max_{1 \leq i \leq d} \frac{\|g_i v\|}{\|v\|} \geq \delta \cdot \min_{1 \leq i \leq d} \|g_i\| \geq e^{-2\epsilon \cdot m} \cdot e^{(\lambda_1 - \epsilon) \cdot m} > e^{(\lambda_1 - 3\epsilon) \cdot m}$$

as claimed. □

*Proof of Proposition 7.2.* We fix a small  $0 < \delta < \frac{\epsilon}{3}$ ; with  $\ell_0$  to be determined later. Lemma 7.6 provides a set  $\mathcal{G} \subset \Gamma^d$  of  $d$ -tuples  $(g_1, \dots, g_d)$  and  $\gamma > 0$  so that

$$\begin{aligned} \left| \frac{1}{m} \log \sigma_j(g_i) - \lambda_j \right| &< \delta \quad (1 \leq i, j \leq d) \\ \text{vol}(\theta(g_1), \dots, \theta(g_d)) &> e^{-\delta \cdot \ell} \\ \text{vol}(\theta(g_1^{\text{tr}}), \dots, \theta(g_d^{\text{tr}})) &> e^{-\delta \cdot \ell} \end{aligned}$$

and for any  $A \subset \mathbb{T}^d$

$$\mu_n(A)^d - e^{-\gamma \cdot \ell} \leq \sum_{\vec{g} \in \mathcal{G}} \nu^{*\ell}(g_1) \cdots \nu^{*\ell}(g_d) \cdot \mu_{n-\ell}(g_1^{-1}A \cap \cdots \cap g_d^{-1}A).$$

We apply this to the set  $A = \text{Nbd}_\rho(X) = \bigcup_{x \in X} \mathbf{B}_{x,\rho}$  of well separated small balls on the torus, and fix a  $d$ -tuple  $(g_1, \dots, g_d) \in \mathcal{G}$  with

$$\begin{aligned} \mu_n(A)^d - e^{-\gamma \cdot \ell} &\leq \mu_{n-\ell}(g_1^{-1}A \cap \cdots \cap g_d^{-1}A) \\ &= \mu_{n-\ell}\left(\bigcup_{x_1, \dots, x_d \in X} g_1^{-1}(\mathbf{B}_{x_1,\rho}) \cap \cdots \cap g_d^{-1}(\mathbf{B}_{x_d,\rho})\right). \end{aligned}$$

Consider the components  $C_{x_1, \dots, x_d} = g_1^{-1}(\mathbf{B}_{x_1,\rho}) \cap \cdots \cap g_d^{-1}(\mathbf{B}_{x_d,\rho})$ , indexed by  $d$ -tuples  $\vec{x} = (x_1, \dots, x_d) \in X^d$ , of the union in the right hand side. We shall show that most of these components are empty; in fact, there are at most  $|X|$ -many ones with  $C_{\vec{x}} \neq \emptyset$ . We shall also show that these non-empty components are  $r'$ -separated and have diameter less than  $\rho'$ . So choosing one point  $y$  from each non-empty component  $C_{\vec{x}}$  of  $g_1^{-1}A \cap \cdots \cap g_d^{-1}A$ , we obtain a set  $Y$  with the desired properties.

Let  $\vec{x} = (x_1, \dots, x_d)$  and  $\vec{x}' = (x'_1, \dots, x'_d)$  be two  $d$ -tuples from  $X$ , where  $C_{\vec{x}}$  and  $C_{\vec{x}'}$  are not empty, and assume that  $x_1 = x'_1 = x$ . Then  $g_1^{-1}(\mathbf{B}_{x,\rho})$  intersects both  $g_j^{-1}(\mathbf{B}_{x_j,\rho})$  and  $g_j^{-1}(\mathbf{B}_{x'_j,\rho})$ . Applying  $g_j$  it follows that the set  $(g_j g_1^{-1})(\mathbf{B}_{x,\rho})$  intersects the  $\rho$ -balls around points  $x_j, x'_j \in X$ , which yields:

$$\|x_j - x'_j\| < 2\rho + \|g_j\| \cdot \|g_1^{-1}\| \cdot \rho < (2 + e^{(\lambda_1 + \delta) \cdot \ell} \cdot e^{(-\lambda_d + \delta) \cdot \ell}) \cdot \rho.$$

By our choice of  $c_3 = \lambda_1 - \lambda_d + 1$  it follows that  $\|x_j - x'_j\| < r$  and therefore  $x_j = x'_j$ . This consideration applies to all  $j = 2, \dots, d$ . So  $\vec{x} = \vec{x}'$ .

Let us choose representatives  $y \in C_{\vec{x}}$  in non-empty components of  $g_1^{-1}A \cap \cdots \cap g_d^{-1}A$  and form the set  $Y$ . We just showed that associating  $x_1$  to  $y \in C_{x_1, \dots, x_d}$  is an *injective* map  $Y \rightarrow X$ , so  $|Y| \leq |X|$ .

Let us show that  $Y$  is  $r'$ -separated. Let  $y \in C_{\vec{x}}$  and  $y' \in C_{\vec{x}'}$  and  $y \neq y'$ . Then  $x_1 \neq x'_1 \in X$ , while  $g_1 y \in \mathbf{B}_{x_1,\rho}$  and  $g_1 y' \in \mathbf{B}_{x'_1,\rho}$ . Therefore

$$r < \|x_1 - x'_1\| \leq 2\rho + \|g_1\| \cdot \|y - y'\|.$$

Since  $\rho$  is much smaller than  $r$ , and  $\|g_1\| < \frac{1}{2} \cdot e^{(\lambda_1 + \delta) \cdot \ell} < e^{(\lambda_1 + \epsilon) \cdot \ell}$ , we have

$$\|y - y'\| > \|g_1\|^{-1} \cdot (r - 2\rho) > e^{-(\lambda_1 + \epsilon) \cdot \ell} \cdot r = r'$$

as claimed.

Let  $C_{\vec{x}}$  be a non-empty component and  $y \in C_{\vec{x}}$ . We claim that  $C_{\vec{x}} \subset \mathbf{B}_{y,\rho'}$ . Indeed, for any  $z \in C_{\vec{x}}$  and every  $i = 1, \dots, d$  both  $g_i y$  and  $g_i z$  are in  $\mathbf{B}_{x_i,\rho}$ , so that

$$\max_{1 \leq i \leq d} \|g_i y - g_i z\| \leq 2\rho.$$

The above distances are measured on the torus. But they are so small that the whole picture may safely be lifted to  $\mathbb{R}^d$ , and one might think of the vector  $v = y - z$

being such that

$$\max_{1 \leq i \leq d} \|g_i v\| \leq 2\rho.$$

By Lemma 7.8 and the geometry of  $g_1, \dots, g_d$  this implies that

$$\|y - z\| = \|v\| < e^{-(\lambda_1 - 3\delta) \cdot \ell} \cdot 2\rho < e^{-(\lambda_1 - \epsilon) \cdot \ell} \cdot \rho = \rho'.$$

Therefore

$$g_1^{-1}A \cap \dots \cap g_d^{-1}A \subset \bigcup_{y \in Y} B_{y, \rho'}$$

and

$$\mu_{n-\ell}(\bigcup_{y \in Y} B_{y, \rho'}) \geq \mu_n(\bigcup_{x \in X} B_{x, \rho})^d - e^{-\gamma \cdot \ell}$$

as required.  $\square$

### 7.C. Rational approximation: proof of Proposition 7.3.

We shall need the following technical lemma, which gives a sufficient condition for a linear combination of  $d$  very proximal elements in  $\mathrm{SL}_d(\mathbb{R})$  to be invertible.

**Lemma 7.9.** *Given  $g_1, \dots, g_d \in \mathrm{SL}_d(\mathbb{R})$  and constants  $c_1, \dots, c_d$  let*

$$\rho = \max_{1 \leq i \leq d} \rho(g_i), \quad C = \max_{1 \leq i, j \leq d} \frac{|c_i|}{|c_j|}, \quad L = \max_{1 \leq i, j \leq d} \frac{\|g_i\|}{\|g_j\|}$$

and let  $v = \min(v_1, v_2)$ , where

$$v_1 = \mathrm{vol}(\theta(g_1), \dots, \theta(g_d)), \quad v_2 = \mathrm{vol}(\theta(g_1^{\mathrm{tr}}), \dots, \theta(g_d^{\mathrm{tr}})).$$

Assume that  $\rho < (20d^2CL)^{-1} \cdot v^3$ . Then the matrix  $h = \sum_{i=1}^d c_i g_i$  is invertible.

*Proof.* The idea is as follows: the transversality parameter  $v_2 > 0$  provides a lower bound on the largest angle between an arbitrary vector  $z$  and the hyperplanes  $H(g_i)$  of "slow growth". This lower bound and the proximality parameter  $\rho$  show that any given vector  $z$  is stretched significantly by at least some of the maps  $g_i$ ; in addition, for these maps  $g_i \bar{z}$  is close to the axis  $\theta(g_i)$ . The fact that these directions are in sufficiently general position (controlled by  $v_1$ ), is used to show that the longer among the images  $g_i z$  do not to cancel each other, and cannot be offset by the shorter images  $g_j z$  either. The details follow.

Given a unit vector  $\|z\| = 1$  reorder the  $g_i$ -s so that

$$\alpha_i = d_{\mathrm{ang}}(\bar{z}, H(g_i)) = d_{\mathrm{ang}}(\bar{z}^\perp, \theta(g_i^{\mathrm{tr}}))$$

decrease:  $\alpha_1 \geq \dots \geq \alpha_d$ . Let  $\beta = 4d\rho/v$  and define  $k = \max\{1 \leq i \leq d : \alpha_i > \beta\}$ . Denoting  $x_i = c_i g_i z$  we shall prove that

$$(7.18) \quad \|x_1 + \dots + x_k\| > \|x_{k+1}\| + \dots + \|x_d\|$$

thereby verifying that  $hz = x_1 + \dots + x_d \neq 0$ . Since  $z$  was an arbitrary unit vector  $h$  is non-singular.

Let  $\bar{y}_i$  denote the projection of  $\theta(g_i^{\text{tr}})$  and to  $\bar{z}^\perp$ . Then  $\text{vol}(\bar{y}_1, \dots, \bar{y}_d) = 0$  and it follows from Lemma 7.7 that

$$\sum_{i=1}^d \alpha_i = \sum_{i=1}^d d_{\text{ang}}(\bar{z}^\perp, \theta(g_i^{\text{tr}})) \geq \frac{v}{\sqrt{2}}.$$

Therefore,  $\alpha_1 \geq v/2d$ , which in turn is bigger than  $\beta = 4d\rho/v$  by the assumptions on  $\rho$ . Hence we are guaranteed that  $k \geq 1$ . Using Lemma 4.1 (3), for  $1 \leq i \leq k$  we have (with  $\bar{x}_i$  denoting the unit vector in direction  $x_i$ )

$$\|x_i\| \geq |c_i| \cdot \|g_i\| \cdot \alpha_i, \quad d_{\text{ang}}(\bar{x}_i, \theta(g_i)) \leq \frac{\rho}{\beta}.$$

Thus applying Lemma 7.7 to

$$t = \text{vol}(\bar{x}_1, \dots, \bar{x}_k, \theta(g_{k+1}), \dots, \theta(g_d)), \quad \text{and} \quad \text{vol}(\theta(g_1), \dots, \theta(g_d)) \geq v$$

gives

$$t > v - \frac{\sqrt{2}d\rho}{\beta} > \frac{v}{2}.$$

Since  $t \leq d_{\text{ang}}(\bar{x}_1, \text{span}(x_2, \dots, x_k))$  it follows that

$$\|x_1 + \dots + x_k\| \geq \|x_1\| \cdot t \geq |c_1| \cdot \|g_1\| \cdot \alpha_1 t \geq |c_1| \cdot \|g_1\| \cdot \frac{v^2}{4d}.$$

At the same time, for  $k < i \leq d$  one has

$$\|x_i\| \leq |c_i| \cdot \|g_i\| \cdot \sqrt{\alpha_i^2 + \rho(g_i)^2} < CL \cdot |c_1| \cdot \|g_1\| \cdot \frac{5d\rho}{v}$$

using  $\alpha_i \leq \beta = 4d\rho/v$ ,  $c_i < Cc_1$ ,  $\|g_i\| \leq L\|g_1\|$ . Hence (7.18) follows from the assumption  $\rho < (20d^2CL)^{-1} \cdot v^3$ .  $\square$

*Proof of Proposition 7.3.* We set  $\epsilon = \frac{\lambda_1 - \lambda_2}{10}$ . Let  $\gamma > 0$  be the corresponding constant from Lemma 7.6, and set  $\delta = \frac{\gamma}{10d^2(\lambda_1 + 2\epsilon)}$ . The constant  $c_4$  is chosen so that for large  $m$  and  $\rho > e^{-c_4 m}$  we have:

$$(7.19) \quad e^{-\gamma m} < \rho^{d\delta}$$

and

$$(7.20) \quad e^{d(\lambda_1 + 2\epsilon) \cdot m} < \rho^{-\frac{1}{10}}.$$

Applying Lemma 7.6 we obtain a set  $\mathcal{G} \subset \Gamma^d$  of  $d$ -tuples  $\vec{g} = (g_1, \dots, g_d)$  with

$$\begin{aligned} \left| \frac{1}{m} \log \|g_i\| - \lambda_1 \right| &< \epsilon, & \left| \frac{1}{m} \log \sigma_2(g_i) - \lambda_2 \right| &< \epsilon, \\ \text{vol}(\theta(g_1), \dots, \theta(g_d)) &> e^{-\epsilon m}, \\ \text{vol}(\theta(g_1^{\text{tr}}), \dots, \theta(g_d^{\text{tr}})) &> e^{-\epsilon m}, \end{aligned}$$

and such that

$$\mu_n(\mathbf{B}_{z,\rho})^d - e^{-\gamma \cdot m} < \sum_{\vec{g} \in \mathcal{G}} \nu^{*m}(g_1) \cdots \nu^{*m}(g_d) \cdot \mu_{n-m} \left( \bigcap_{i=1}^d g_i^{-1}(\mathbf{B}_{z,\rho}) \right).$$



The assumption  $\mu_n(\mathbf{B}_{z,\rho}) > \rho^\delta$  implies (7.19) that

$$\mu_n(\mathbf{B}_{z,\rho})^d > \rho^{d\delta} > e^{-\gamma m}.$$

In particular, there exists a  $d$ -tuple  $(g_1, \dots, g_d) \in \mathcal{G}$  with  $\mu_{n-m}(g_1^{-1}(\mathbf{B}_{z,\rho}) \cap \dots \cap g_d^{-1}(\mathbf{B}_{z,\rho})) > 0$ . In particular, there exists  $w \in \mathbb{T}^d$  such that

$$\{g_1 w, g_2 w, \dots, g_d w\} \subset \mathbf{B}_{z,\rho}.$$

Choose  $\vec{w}, \vec{z} \in \mathbb{R}^d$  projecting to  $w, z \in \mathbb{T}^d$ . Then for some integer vectors  $\vec{a}_i \in \mathbb{Z}^d$

$$(7.21) \quad \|g_i \vec{w} - \vec{a}_i - \vec{z}\| < \rho \quad (i = 1, \dots, d).$$

Let  $c_1 = \dots = c_{d-1} = 1, c_d = 1 - d$ , so that  $\sum c_i = 0$  and  $\sum |c_i| < 2d$ . Combining the inequalities (7.21) with coefficients  $c_i$  we get

$$\|h \vec{w} - \vec{b}\| < 2d\rho$$

where  $h = \sum_{i=1}^d c_i g_i$  is an integer  $d \times d$  matrix and  $\vec{b} = \sum_{i=1}^d c_i \vec{a}_i$  is an integer vector.

Our choice of  $\epsilon = \frac{\lambda_1 - \lambda_2}{10}$  and the following properties of  $g_1, \dots, g_d$

$$\begin{aligned} \max \rho(g_i) &< \frac{e^{(\lambda_2 + \epsilon) \cdot m}}{e^{(\lambda_1 - \epsilon) \cdot m}}, \quad \max \frac{|c_i|}{|c_j|} < 2d < e^{\epsilon \cdot m}, \quad \max \frac{\|g_i\|}{\|g_j\|} < e^{2\epsilon \cdot m}, \\ \text{vol}(\theta(g_1), \dots, \theta(g_d)) &> e^{-\epsilon \cdot m}, \quad \text{vol}(\theta(g_1^{\text{tr}}), \dots, \theta(g_d^{\text{tr}})) > e^{-\epsilon \cdot m}, \end{aligned}$$

imply that the assumptions of Lemma 7.9 are satisfied. Thus the integer matrix  $h$  is invertible, and its determinant  $q = \det(h)$  is a non-zero integer; in particular  $|q| \geq 1$ . Let  $k \in M_{d \times d}(\mathbb{Z})$  be such that  $hk = q \cdot I_d$ , and set  $\vec{p} = g_1 k \vec{b} \in \mathbb{Z}^d$ . We have

$$(7.22) \quad \left\| \vec{z} - \frac{\vec{p}}{q} \right\| \leq \|g_1 h^{-1}\| \cdot \|h \vec{w} - \vec{b}\| + \rho < (1 + 2d \|g_1\| \|h^{-1}\|) \cdot \rho.$$

Let us estimate  $|q|$  and  $\|h^{-1}\|$  using  $\|g_i\| < e^{(\lambda_1 + \epsilon) \cdot m}$ . For  $m > m(\epsilon)$  we may also assume that  $\rho < e^{-c_4 \cdot m}$ .

$$\begin{aligned} \|h\| &\leq \sum_{i=1}^d |c_i| \cdot \|g_i\| \leq 2de^{(\lambda_1 + \epsilon) \cdot m}, \\ |q| = |\det(h)| &\leq \|h\|^d < (2de^{(\lambda_1 + \epsilon) \cdot m})^d < \rho^{-\frac{1}{10}}, \end{aligned}$$

Since  $|q| \geq 1$  we also have

$$\|h^{-1}\| = |q|^{-1} \cdot \|k\| \leq \|k\| \leq \text{const}_d \cdot \|h\|^{d-1},$$

and using (7.20) we deduce that

$$1 + 2d \|g_1\| \cdot \|h^{-1}\| < \text{const}'_d \cdot (2de^{(\lambda_1 + \epsilon) \cdot m})^d < \rho^{-\frac{1}{10}}.$$

Combining this estimate with (7.22) gives

$$\left\| \vec{z} - \frac{\vec{p}}{q} \right\| < \rho^{\frac{9}{10}}, \quad |q| < \rho^{-\frac{1}{10}}$$

as claimed.  $\square$

### 7.D. Final Bootstrap: proof of Proposition 7.4.

#### Lemma 7.10.

Given  $\epsilon_1, \epsilon_2 > 0$  there exist  $\delta > 0$  and  $m_0$  so that for  $m \geq m_0$  any set  $\mathcal{G} \subset \Gamma$  with  $\nu^{*m}(\mathcal{G}) > e^{-\delta \cdot m}$  contains a subset  $\mathcal{F} \subset \mathcal{G}$  with cardinality  $|\mathcal{F}| > e^{\delta \cdot m}$ , such that

$$\left| \frac{1}{m} \log \sigma_j(g) - \lambda_j \right| < \epsilon_1 \quad (j = 1, \dots, d, \quad g \in \mathcal{F})$$

and every  $d$ -element subset  $\{g_1, \dots, g_d\} \subset \mathcal{F}$  satisfies

$$\text{vol}(\theta(g_1^{\text{tr}}), \dots, \theta(g_d^{\text{tr}})) > e^{-\epsilon_2 \cdot m}.$$

*Proof.* Let  $\mathcal{G}_{\text{len}} = \{g \in \Gamma : |\frac{1}{m} \log \sigma_j(g) - \lambda_j| < \epsilon_1, (1 \leq j \leq d)\}$ . By Theorem 4.3 there exist  $N = N(\epsilon_1)$  and  $c_1 > 0$  so that for  $m > N$  the set

$$\mathcal{G}_{\text{len}} = \left\{ g \in \Gamma : \left| \frac{1}{m} \log \sigma_j(g) - \lambda_j \right| < \epsilon_1 \quad (j = 1, 2) \right\}$$

has  $\nu^{*m}(\mathcal{G}_{\text{len}}) > 1 - e^{-c_1 \cdot m}$ . By Theorem 4.4 given  $\epsilon_2 > 0$  there is  $c_2 > 0$  so that for any hyperplane  $H$ :

$$\nu^{*m} \{g \in \Gamma : d_{\text{ang}}(\theta(g^{\text{tr}}), H) \leq e^{-\epsilon_2 \cdot m}\} < e^{-c_2 \cdot m}.$$

Let us take positive  $\delta < \min(c_1, c_2/d)$ . For such  $\delta$  and large  $m$

$$(7.23) \quad e^{-\delta \cdot m} - e^{-c_1 \cdot m} - (e^{\delta \cdot m})^{d-1} \cdot e^{-c_2 \cdot m} > 0.$$

Let  $\mathcal{G}$  with  $\nu^{*m}(\mathcal{G}) > e^{\delta \cdot m}$  be given. We shall form the subset  $\mathcal{F} \subset \mathcal{G}$  by choosing inductively elements from  $\mathcal{G}' = \mathcal{G} \cap \mathcal{G}_{\text{len}}$  by induction. Suppose  $g_1, \dots, g_n$  are already chosen. For the next element  $g_{n+1}$  we can choose any  $g \in \mathcal{G}'$  for which the axis  $\theta(g^{\text{tr}})$  makes angle of at least  $e^{-\epsilon_2 \cdot m}$  with all hyperplanes of the form

$$\theta(g_{i_1}^{\text{tr}}) \oplus \dots \oplus \theta(g_{i_{d-1}}^{\text{tr}})$$

where  $i_1 < \dots < i_{d-1}$  is a  $(d-1)$ -element subset of  $\{1, \dots, n\}$ . There are less than  $n^{d-1}$  such hyperplanes. It follows that

$$\begin{aligned} & \nu^{*m} \left( \mathcal{G}' \setminus \bigcup_{1 \leq i_1 < \dots < i_{d-1} \leq n} \left\{ g : d_{\text{ang}} \left( \theta(g^{\text{tr}}), \theta(g_{i_1}^{\text{tr}}) \oplus \dots \oplus \theta(g_{i_{d-1}}^{\text{tr}}) \right) < e^{-\epsilon_2 \cdot m} \right\} \right) \\ & > e^{-\delta \cdot m} - e^{-c_1 \cdot m} - n^{d-1} \cdot e^{-c_2 \cdot m}, \end{aligned}$$

and in view of (7.23) the RHS is positive as long as  $n \leq [e^{\delta \cdot m}]$ . This allows to construct the desired set  $\mathcal{F}$  with at least  $e^{\delta \cdot m}$  elements.  $\square$

*Proof of Proposition 7.4.*

Let  $\delta > 0$  be associated to  $\epsilon_1 = \epsilon_2 = \frac{\epsilon}{3}$  in Lemma 7.10, take  $\omega = \delta/2$  and  $m_0$  be large enough. The basic relation

$$\mu_n(\mathbb{W}_{Q,r}) = \sum_{g \in \Gamma} \nu^{*m}(g) \cdot \mu_{n-m}(g^{-1}(\mathbb{W}_{Q,r}))$$

implies that the set

$$\mathcal{G} = \left\{ g \in \Gamma : \mu_{n-m}(g^{-1}(W_{Q,r})) > \mu_n(W_{Q,r}) - e^{-\delta \cdot m} \right\}$$

has  $\nu^{*m}(\mathcal{G}) > e^{-\delta \cdot m}$ . Let  $\mathcal{F} \subset \mathcal{G}$  be a subset of size  $|\mathcal{F}| > e^{\delta \cdot m}$  of well shaped elements in general position provided by Lemma 7.10. We shall consider the possible intersections of the sets

$$g^{-1}(W_{Q,r}) = \bigcup_{x \in R_Q} g^{-1}(B_{x,r}) \quad (g \in \mathcal{F}).$$

Note that the set  $R_Q$  of centers of the  $r$ -balls which form  $W_{Q,r}$  is  $Q^{-2}$ -separated:

$$\left\| \left( \frac{p_1}{q}, \dots, \frac{p_d}{q} \right) - \left( \frac{p'_1}{q'}, \dots, \frac{p'_d}{q'} \right) \right\| = \left\| \left( \frac{q'p_1 - qp'_1}{q \cdot q'}, \dots, \frac{q'p_d - qp'_d}{q \cdot q'} \right) \right\| \geq \frac{1}{qq'} \geq \frac{1}{Q^2}.$$

Suppose that for  $x, y \in R_Q$  and  $g, h \in \mathcal{F}$  the ellipses  $g^{-1}(B_{x,r})$  and  $h^{-1}(B_{y,r})$  have a common point, say  $w$ . We have  $\|x - gw\| < r$ ,  $\|y - hw\| < r$ , and  $\|g^{-1}\|, \|h^{-1}\| < e^{(-\lambda_d + \epsilon_1) \cdot m}$ . Note also that  $-\lambda_d < d\lambda_1$ , and we may assume that  $2e^{-(\lambda_d + \epsilon_1) \cdot m} < e^{d\lambda_1 \cdot m}$ . Therefore

$$\begin{aligned} \|g^{-1}x - h^{-1}y\| &\leq \|g^{-1}x - w\| + \|w - h^{-1}y\| \\ &< \|g^{-1}\| \cdot \|x - gw\| + \|h^{-1}\| \cdot \|y - hw\| \\ &< 2e^{(-\lambda_d + \epsilon_1) \cdot m} \cdot r < e^{d\lambda_1 \cdot m} \cdot r < \frac{1}{Q^2}. \end{aligned}$$

Since  $g^{-1}x$  and  $h^{-1}y$  belong to the  $Q^{-2}$ -separated set  $R_Q$ , they coincide:  $g^{-1}x = h^{-1}y = z \in R_Q$ .

This computation shows that for any  $d$ -element subset  $\{g_1, \dots, g_d\} \subset \mathcal{F}$  we have

$$\bigcap_{i=1}^d g_i^{-1}(W_{Q,r}) = \bigcup_{z \in R_Q} \left( \bigcap_{i=1}^d g_i^{-1}(B_{g_i z, r}) \right).$$

The conditions on  $\mathcal{F}$  show, using Lemma 7.8, that for any  $d$ -element subset  $\{g_1, \dots, g_d\} \subset \mathcal{F}$  and every  $v \in \mathbb{R}^d$

$$\max_{1 \leq i \leq d} \|g_i v\| \geq e^{(\lambda_1 - \epsilon) \cdot m} \cdot \|v\|.$$

This implies that on the torus  $\mathbb{T}^d$ :

$$\bigcap_{i=1}^d g_i^{-1}(B_{g_i z, r}) \subset B_{z, e^{-(\lambda_1 - \epsilon) \cdot m} \cdot r}$$

and therefore for any  $d$ -element subset  $\{g_1, \dots, g_d\} \subset \mathcal{F}$  we have

$$\bigcap_{i=1}^d g_i^{-1}(W_{Q,r}) \subset W_{Q, e^{-(\lambda_1 - \epsilon) \cdot m} \cdot r}$$

For  $g \in \mathcal{F}$  let  $E_g = g^{-1}(W_{Q,r}) \setminus W_{Q,e^{-(\lambda_1 - \epsilon) \cdot m}, r}$ . We just showed that the collection  $\{E_g \mid g \in \mathcal{F}\}$  has no  $d$ -fold intersections. Thus

$$d > \int \sum_{g \in \mathcal{F}} 1_{E_g}(x) d\mu_{n-m}(x) = \sum_{g \in \mathcal{F}} \mu_{n-m}(E_g).$$

Thus for at least one  $h \in \mathcal{F} \subset \mathcal{G}$  one has

$$\mu_{n-m}(E_h) \leq \frac{d}{|\mathcal{F}|} < d \cdot e^{-\delta \cdot m}.$$

Therefore:

$$\begin{aligned} \mu_{n-m}(W_{Q,e^{-(\lambda_1 - \epsilon) \cdot m}, r}) &= \mu_{n-m}(h^{-1}(W_{Q,r})) - \mu_{n-m}(E_h) \\ &> \mu_n(W_{Q,r}) - e^{-\delta \cdot m} - d \cdot e^{-\delta \cdot m} \\ &> \mu_n(W_{Q,r}) - e^{-\omega \cdot m}, \end{aligned}$$

assuming  $m > m_0$  where  $m_0$  is large enough.  $\square$

#### REFERENCES

- [1] D. Berend, *Multi-invariant sets on compact abelian groups*, Trans. Amer. Math. Soc. **286** (1984), no. 2, 505–535.
- [2] Y. Benoist and F. Quint, *Mesures stationnaires et fermés invariants des espaces homogènes*.
- [3] P. Bougerol and J. Lacroix, *Products of random matrices with applications to Schrödinger operators*, Progress in Probability and Statistics, vol. 8, Birkhäuser Boston Inc., Boston, MA, 1985.
- [4] J. Bourgain, *On the Erdős-Volkmann and Katz-Tao ring conjectures*, Geom. Funct. Anal. **13** (2003), no. 2, 334–365.
- [5] ———, *The discretized sum product and projection theorems* (2009).
- [6] J. Bourgain, A. Furman, E. Lindenstrauss, and S. Mozes, *Invariant measures and stiffness for non-abelian groups of toral automorphisms*, C. R. Math. Acad. Sci. Paris **344** (2007), no. 12, 737–742 (English, with English and French summaries).
- [7] J. Bourgain and A. Gamburd, *On the spectral gap for finitely-generated subgroups of  $SU(2)$* , Invent. Math. **171** (2008), no. 1, 83–121.
- [8] J. Bourgain, A. Gamburd, and P. Sarnak, *Sieving and expanders*, C. R. Math. Acad. Sci. Paris **343** (2006), no. 3, 155–159 (English, with English and French summaries).
- [9] M. Burger, *Kazhdan constants for  $SL(3, \mathbf{Z})$* , J. Reine Angew. Math. **413** (1991), 36–67.
- [10] M. Einsiedler and E. Lindenstrauss, *Rigidity properties of  $\mathbb{Z}^d$ -actions on tori and solenoids*, Electron. Res. Announc. Amer. Math. Soc. **9** (2003), 99–110 (electronic).
- [11] H. Furstenberg, *Disjointness in ergodic theory, minimal sets, and a problem in Diophantine approximation*, Math. Systems Theory **1** (1967), 1–49.
- [12] ———, *Stiffness of group actions*, Lie groups and ergodic theory (Mumbai, 1996), Tata Inst. Fund. Res. Stud. Math., vol. 14, Tata Inst. Fund. Res., Bombay, 1998, pp. 105–117.
- [13] K. J. Falconer, *Hausdorff dimension and the exceptional set of projections*, Mathematika **29** (1982), no. 1, 109–115. MR673510 (83m:28014)
- [14] Harry Furstenberg, *Noncommuting random products*, Trans. Amer. Math. Soc. **108** (1963), 377–428. MR0163345 (29 #648)
- [15] H. Furstenberg and Y. Kifer, *Random matrix products and measures on projective spaces*, Israel J. Math. **46** (1983), no. 1-2, 12–32.
- [16] I. Ya. Gol'dsheid and G. A. Margulis, *Lyapunov exponents of a product of random matrices*, Uspekhi Mat. Nauk **44** (1989), no. 5(269), 13–60 (Russian); English transl., Russian Math. Surveys **44** (1989), no. 5, 11–71.

- [17] Y. Guivarc'h and A. Raugi, *Products of random matrices: convergence theorems*, Random matrices and their applications (Brunswick, Maine, 1984), 1986, pp. 31–54.
- [18] ———, *Propriétés de contraction d'un semi-groupe de matrices inversibles. Coefficients de Liapunoff d'un produit de matrices aléatoires indépendantes*, Israel J. Math. **65** (1989), no. 2, 165–196.
- [19] Y. Guivarc'h and A. N. Starkov, *Orbits of linear group actions, random walks on homogeneous spaces and toral automorphisms*, Ergodic Theory Dynam. Systems **24** (2004), no. 3, 767–802.
- [20] B. Kalinin and A. Katok, *Invariant measures for actions of higher rank abelian groups*, Smooth ergodic theory and its applications (Seattle, WA, 1999), Proc. Sympos. Pure Math., vol. 69, Amer. Math. Soc., Providence, RI, 2001, pp. 593–637.
- [21] A. Katok and R. J. Spatzier, *Invariant measures for higher-rank hyperbolic abelian actions*, Ergodic Theory Dynam. Systems **16** (1996), no. 4, 751–778.
- [22] N. Katz and T. Tao, *Some connections between Falconer's distance set conjecture and sets of Furstenberg type*, New York J. Math. **7** (2001), 149–187 (electronic).
- [23] É. Le Page, *Théorèmes limites pour les produits de matrices aléatoires*, Probability measures on groups (Oberwolfach, 1981), Lecture Notes in Math., vol. 928, Springer, Berlin, 1982, pp. 258–303 (French).
- [24] G. A. Margulis, *Problems and conjectures in rigidity theory*, Mathematics: frontiers and perspectives, Amer. Math. Soc., Providence, RI, 2000, pp. 161–174.
- [25] Pertti Mattila, *Geometry of sets and measures in Euclidean spaces*, Cambridge Studies in Advanced Mathematics, vol. 44, Cambridge University Press, Cambridge, 1995. Fractals and rectifiability. MR1333890 (96h:28006)
- [26] R. Muchnik, *Semigroup actions on  $\mathbb{T}^n$* , Geom. Dedicata **110** (2005), 1–47.
- [27] Yuval Peres and Wilhelm Schlag, *Smoothness of projections, Bernoulli convolutions, and the dimension of exceptions*, Duke Math. J. **102** (2000), no. 2, 193–251. MR1749437 (2001d:42013)
- [28] M. Ratner, *Interactions between ergodic theory, Lie groups, and number theory*, Proceedings of the international congress of mathematicians, vol. 1, 2 (Zürich, 1994), 1995, pp. 157–182.
- [29] D. J. Rudolph,  *$\times 2$  and  $\times 3$  invariant measures and entropy*, Ergodic Theory Dynam. Systems **10** (1990), no. 2, 395–406.

INSTITUTE FOR ADVANCED STUDY, PRINCETON, NJ 08540, USA

UNIVERSITY OF ILLINOIS AT CHICAGO, IL 60607, USA

PRINCETON UNIVERSITY, PRINCETON, NJ 08544, USA and THE HEBREW UNIVERSITY, 91904 JERUSALEM, ISRAEL

THE HEBREW UNIVERSITY, 91904 JERUSALEM, ISRAEL